

=> fil hcaplus
FILE 'HCAPLUS' ENTERED AT 15:31:14 ON 04 MAR 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)

Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December 26, 1996), unless otherwise indicated in the original publications. The CA Lexicon is the copyrighted intellectual property of the the American Chemical Society and is provided to assist you in searching databases on STN. Any dissemination, distribution, copying, or storing of this information, without the prior written consent of CAS, is strictly prohibited.

FILE COVERS 1907 - 4 Mar 2003 VOL 138 ISS 10
FILE LAST UPDATED: 3 Mar 2003 (20030303/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

=> d que 120
L1 8 SEA FILE=REGISTRY ADSDINIKTGTTDIGSNTTVKTGDLVTYDKENGMHKKVFYSFIDD
KNHNKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDYYPNSID
TK.Y.S.L.Y.F.G.V.G.D.G.I.G.I.A.V.I.H.L.Y.QPDFKTILESPTDKKVGWKVIF
NNMVNQNWGPYDRDSWNPVYGNQLFMKTRNGSMKAADNF/SQSP
L2 6 SEA FILE=REGISTRY ADSDINIKTGTTDIGSNTTVKTGDLVTYDKENGMHKKVFYSFIDD
KNHNKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDYYPNSID
T.E.M.T.T.G.N.N.T.D.T.K.G.L.G.N.S.G.T.K.VQPDFKTILESPTDKKVGWKVIF
NNMVNQNWGPYDRDSWNPVYGNQLFMKTRNGSMKAADNF/SQSP
L3 8 SEA FILE=REGISTRY L1 OR L2
L4 3 SEA FILE=HCAPLUS L3
L5 1 SEA FILE=HCAPLUS L4 AND PORE#
L6 1 SEA FILE=HCAPLUS L4 AND STEM#
L7 1 SEA FILE=HCAPLUS L4 AND LUMEN?
L8 1 SEA FILE=HCAPLUS L4 AND TRANSMEMBRANE(3A)CHANNEL?
L9 0 SEA FILE=HCAPLUS L4 AND HEPTAMER?
L10 1 SEA FILE=HCAPLUS L4 AND HETEROLOG?
L11 1 SEA FILE=HCAPLUS L4 AND (SENSOR? OR BIOSENSOR?)
L12 1 SEA FILE=HCAPLUS L4 AND METAL
L13 0 SEA FILE=HCAPLUS L4 AND CURRENT?
L14 0 SEA FILE=HCAPLUS L4 AND DIGITAL
L20 1 SEA FILE=HCAPLUS (L5 OR L6 OR L7 OR L8 OR L9 OR L10 OR L11 OR
L12 OR L13 OR L14)

=> d ibib abs 120 1

L20 ANSWER 1 OF 1 HCAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1999:96264 HCAPLUS
DOCUMENT NUMBER: 130:165159
TITLE: Designed staphylococcal hemolysin protein
pores as components for metal
biosensors
INVENTOR(S): Bayley, Hagan; Braha, Orit; Kasianowicz, John; Gouaux,
Eric

THIS PAGE BLANK (USPTO)

PATENT ASSIGNEE(S): University of Massachusetts, USA
 SOURCE: PCT Int. Appl., 51 pp.
 CODEN: PIXXD2
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9905167	A1	19990204	WO 1998-US15354	19980724
W: AU, CA, JP, KR				
RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
AU 9885862	A1	19990216	AU 1998-85862	19980724
PRIORITY APPLN. INFO.:			US 1997-53737P	P 19970725
			WO 1998-US15354	W 19980724

AB This invention features a mutant staphylococcal alpha hemolysin (.alpha.HL) polypeptide contg. a **heterologous metal** -binding amino acid. The polypeptide assembles into a heteroheptameric **pore** assembly in the presence of a wild type .alpha.HL polypeptide. Preferably, the **metal**-binding amino acid occupies a position in a **transmembrane channel** of the heteroheptameric **pore** assembly, e.g., an amino acid in the **stem** domain of WT .alpha.HL is substituted with a **heterologous metal**-binding amino acid. More preferably, the **metal**-binding amino acid projects into the **lumen** of the **transmembrane channel**.

REFERENCE COUNT: 1 THERE ARE 1 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

=> d que 122

L1	8	SEA FILE=REGISTRY ADSDINIKTGTTDIGSNTTVKTGDLVTYDKENGMHKKVFYFSFIDD KNHNKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDIYPRNSID TK.Y.S.L.Y.F.G.V.G.D.G.I.G.I.A.V.I.H.L.Y.QPDFKFILESPTDKKVGWKVIF NNMVNQNWGPYDRDSWNPVYGNQLFMKTRNGSMKAADNF/SQSP
L2	6	SEA FILE=REGISTRY ADSDINIKTGTTDIGSNTTVKTGDLVTYDKENGMHKKVFYFSFIDD KNHNKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDIYPRNSID T.E.M.T.T.G.N.N.T.D.T.K.G.L.G.N.S.G.T.K.VQPDFKFILESPTDKKVGWKVIF NNMVNQNWGPYDRDSWNPVYGNQLFMKTRNGSMKAADNF/SQSP
L3	8	SEA FILE=REGISTRY L1 OR L2
L4	3	SEA FILE=HCAPLUS L3
L5	1	SEA FILE=HCAPLUS L4 AND PORE#
L6	1	SEA FILE=HCAPLUS L4 AND STEM#
L7	1	SEA FILE=HCAPLUS L4 AND LUMEN?
L8	1	SEA FILE=HCAPLUS L4 AND TRANSMEMBRANE (3A) CHANNEL?
L9	0	SEA FILE=HCAPLUS L4 AND HEPTAMER?
L10	1	SEA FILE=HCAPLUS L4 AND HETEROLOG?
L11	1	SEA FILE=HCAPLUS L4 AND (SENSOR? OR BIOSENSOR?)
L12	1	SEA FILE=HCAPLUS L4 AND METAL
L13	0	SEA FILE=HCAPLUS L4 AND CURRENT?
L14	0	SEA FILE=HCAPLUS L4 AND DIGITAL
L15	640	SEA FILE=HCAPLUS STAPH? (3A) (HEMOLYSIN# OR HAEMOLYSIN#)
L16	6	SEA FILE=HCAPLUS L15 AND ((SUBSTITUT? OR HETEROLOG? OR REPLAC?) (3A) (AMINO(A)ACID#))
L17	3	SEA FILE=HCAPLUS (L16 AND (PORE# OR STEM# OR LUMEN? OR TRANSMEMBRANE (3A) CHANNEL?))
L18	2	SEA FILE=HCAPLUS L16 AND (SENSOR? OR BIOSENSOR?)

THIS PAGE BLANK (USPTO)

L19 2 SEA FILE=HCAPLUS L16 AND (METAL# OR CURRENT# OR DIGITAL)
L20 1 SEA FILE=HCAPLUS (L5 OR L6 OR L7 OR L8 OR L9 OR L10 OR L11 OR
L12 OR L13 OR L14)
L21 3 SEA FILE=HCAPLUS (L17 OR L18 OR L19)
L22 2 SEA FILE=HCAPLUS L21 NOT L20

=> d ibib abs l22 1-2

L22 ANSWER 1 OF 2 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1997:331282 HCAPLUS

DOCUMENT NUMBER: 127:62190

TITLE: A predicted .beta.-sheet from class S components of
staphylococcal .gamma.-hemolysin is
essential for the secondary interaction of the class F
component

AUTHOR(S): Meunier, O.; Ferreras, M.; Supersac, G.; Hoepfer, F.;
Baba-Moussa, L.; Monteil, H.; Colin, D. A.;
Menestrina, G.; Prevost, G.

CORPORATE SOURCE: Institut de Bacteriologie de la Faculte de Medecine de
Strasbourg, 3, rue Koeberle, Strasbourg, F-67000, Fr.

SOURCE: Biochimica et Biophysica Acta (1997), 1326(2), 275-286
CODEN: BBACAQ; ISSN: 0006-3002

PUBLISHER: Elsevier

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Site-directed mutagenesis was performed on genes encoding HlgA and HlgC,
two of the three proteins expressed from the **staphylococcal**
.gamma.-hemolysin locus, which originate two **pore**
-forming toxins (HlgA+HlgB, HlgC+HlgB). As related proteins, HlgA and
HlgC were found to bind first to cell membranes. **Amino**
acid substitutions concerning residues that would
predictably disrupt a 13 amino acid conserved .beta.-sheet of the Chou and
Fasman secondary structure prediction were created. The mutation of a
threonine into an aspartic acid residue from HlgA (T28D) and from HlgC
(T30D) that would break this predicted N-terminal structure lowered
dramatically the biol. activities towards purely lipidic vesicles,
erythrocytes, and polymorphonuclear cells. The change in secondary
structure was confirmed by Fourier Transform IR spectroscopy. The binding
of mutated and native proteins at the same kind of sites onto
polymorphonuclear cells was evidenced with flow cytometry and
fluorescein-labeled anti-class S antibodies or wild type HlgA or HlgC.
However, the subsequent binding of fluorescein-labeled HlgB to
membrane-bound mutated HlgA or HlgC complexes was inhibited. In
conclusion, the initial binding of class S components is essential for the
subsequent binding of class F components, and a predicted .beta.-sheet
seems to be at least one of the functional domains involved.

L22 ANSWER 2 OF 2 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1994:529301 HCAPLUS

DOCUMENT NUMBER: 121:129301

TITLE: A **pore**-forming protein with a **metal**
-actuated switch

AUTHOR(S): Walker, Barbara; Kasianowicz, John; Krishnasastri,
Musti; Bayley, Hagan

CORPORATE SOURCE: Worcester Found. Exp. Biol., Shrewsbury, MA, 01545,
USA

SOURCE: Protein Engineering (1994), 7(5), 655-62
CODEN: PRENE9; ISSN: 0269-2139

THIS PAGE BLANK (USPTO)

DOCUMENT TYPE: Journal
 LANGUAGE: English

AB **Staphylococcal .alpha.-hemolysin, a pore**
 -forming exotoxin, is a polypeptide of 293 amino acids that is secreted by *Staphylococcus aureus* as a water-sol. monomer. It assembles to form hexameric **pores** in lipid bilayers. Previous studies of **pore** formation have established the involvement of a central glycine-rich loop. Here, the authors show that when five consecutive histidine residues **replace amino acids** 130-134 at the midpoint of the loop, they provide a switch with which **pore** activity can be (i) turned off by micromolar concns. of divalent zinc ions and (ii) turned back on with the chelating agent EDTA. Planar bilayer recordings show that Zn²⁺ and EDTA can act on open channels from either side of the bilayer and thus demonstrate that the central loop lines part of the conductive pathway. The authors' results suggest that genetically-engineered **pore**-forming proteins might make useful components of **metal ion sensors**.

=> d que 123

```
L1      8 SEA FILE=REGISTRY ADSDINIKTGTTDIGSNTTVKTDGLVITYDKENGMHKKVFYSFIDD
        KNHNKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDYYPRNSID
        TK.Y.S.L.Y.F.G.V.G.D.G.I.G.I.A.V.I.H.L.Y.QPDFKFILESPTDKKVGWKVIF
        NNMVNQNWGPYDRDSWNPVYGNQLFMKTRNGSMKAADNF/SQSP
L2      6 SEA FILE=REGISTRY ADSDINIKTGTTDIGSNTTVKTDGLVITYDKENGMHKKVFYSFIDD
        KNHNKLLVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVQLQLPDNEVAQISDYYPRNSID
        T.E.M.T.T.G.N.N.T.D.T.K.G.L.G.N.S.G.T.K.VQPDFKFILESPTDKKVGWKVIF
        NNMVNQNWGPYDRDSWNPVYGNQLFMKTRNGSMKAADNF/SQSP
L3      8 SEA FILE=REGISTRY L1 OR L2
L4      3 SEA FILE=HCAPLUS L3
L5      1 SEA FILE=HCAPLUS L4 AND PORE#
L6      1 SEA FILE=HCAPLUS L4 AND STEM#
L7      1 SEA FILE=HCAPLUS L4 AND LUMEN?
L8      1 SEA FILE=HCAPLUS L4 AND TRANSMEMBRANE (3A) CHANNEL?
L9      0 SEA FILE=HCAPLUS L4 AND HEPTAMER?
L10     1 SEA FILE=HCAPLUS L4 AND HETEROLOG?
L11     1 SEA FILE=HCAPLUS L4 AND (SENSOR? OR BIOSENSOR?)
L12     1 SEA FILE=HCAPLUS L4 AND METAL
L13     0 SEA FILE=HCAPLUS L4 AND CURRENT?
L14     0 SEA FILE=HCAPLUS L4 AND DIGITAL
L20     1 SEA FILE=HCAPLUS (L5 OR L6 OR L7 OR L8 OR L9 OR L10 OR L11 OR
        L12 OR L13 OR L14)
L23     2 SEA FILE=HCAPLUS L4 NOT L20
```

=> d ibib abs 123 1-2

L23 ANSWER 1 OF 2 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2002:906293 HCAPLUS

DOCUMENT NUMBER: 138:8311

TITLE: *Staphylococcus aureus* proteins and nucleic acids and
 their diagnostic and therapeutic uses for
 staphylococcal infections

INVENTOR(S): Masignani, Vega; Mora, Marirosa; Scarselli, Maria

PATENT ASSIGNEE(S): Chiron Spa, Italy

SOURCE: PCT Int. Appl., 49 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

THIS PAGE BLANK (USPTO)

FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002094868	A2	20021128	WO 2002-IB2637	20020327
<p>W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM</p> <p>RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG</p>				
PRIORITY APPLN. INFO.:			GB 2001-7661	A 20010327
<p>AB The invention provides 2821 nucleic acid coding sequences from Staphylococcus aureus strain NCTC 8325 along with their inferred translation products. The proteins are useful for vaccines, immunogenic compns., diagnostics, enzymic studies, and also as targets for antibiotics.</p>				

L23 ANSWER 2 OF 2 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1995:883755 HCAPLUS
DOCUMENT NUMBER: 124:23160
TITLE: Activation of .alpha.-toxin translation in Staphylococcus aureus by the trans-encoded antisense RNA, RNAIII
AUTHOR(S): Morfedt, Eva; Taylor, David; von Gabain, Alexander; Arvidson, Staffan
CORPORATE SOURCE: Microbiol. Tumoriol. Cent., Karolinska Inst., Stockholm, S-171 77, Swed.
SOURCE: EMBO Journal (1995), 14(18), 4569-77
CODEN: EMJODG; ISSN: 0261-4189
PUBLISHER: Oxford University Press
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The synthesis of virulence factors in Staphylococcus aureus is controlled by a regulatory RNA mol., RNAIII, encoded by the agr locus. Transcription of genes coding for secreted toxins and enzymes is stimulated, while transcription of cell-surface protein genes is repressed by RNAIII. In the case of staphylococcal .alpha.-toxin, RNAIII also seems to stimulate translation by an independent mechanism. In this report we show that in a mutant lacking RNAIII the rate of .alpha.-toxin (hla) prodn. relative to the cellular concn. of hla mRNA was reduced 10-fold as compared with the wild-type strain. A 75% complementarity between the 5' end of RNAIII and the 5' untranslated region of the hla transcript suggests a direct interaction between the RNAs. A complex of RNAIII and hla mRNA was demonstrated in exts. of total RNA from the wild-type strain, and also with in vitro synthesized RNAs. RNase T1 digestion expts. revealed that the ribosome binding site of the hla transcript is blocked by intramol. base-pairing. Hybridization with RNAIII prevents this intramol. base-pairing and makes the hla mRNA accessible for translation initiation. This is, to our knowledge, the first example of an antisense RNA that stimulates translation of the target mRNA.

THIS PAGE BLANK (08PT0)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: March 4, 2003, 10:38:18 ; Search time 33 Seconds
(without alignments)
1183.104 Million cell updates/sec

Title: CLAIM6
Perfect score: 1485
Sequence: 1 ADSINIKTGTIGSNTTV.....WTDSSERYKIDWEKEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1485	100.0	293	20	AAV01365
2	424.5	28.6	330	22	AAU33797
3	424.5	28.6	333	22	AAU33797
4	404	27.2	325	22	AAU334287
5	404	27.2	325	22	AAU334287
6	404	27.2	325	22	AAU371100
7	404	27.2	325	22	AAU371186
8	404	27.2	325	22	AAU37477
9	403.5	27.2	327	22	AAU34405
10	403.5	27.2	327	22	AAU371136

11	403.5	27.2	327	22	AAU37564
12	397	26.7	323	22	AAU34395
13	397	26.7	338	22	AAU37509
14	390	26.3	323	14	AAU35084
15	252	17.0	321	22	AAU34275
16	252	17.0	321	22	AAU37101
17	252	17.0	321	22	AAU37128
18	238	16.0	315	14	AAU35083
19	234	15.8	312	22	AAU33771
20	126	8.5	40	20	AAU01366
21	124.5	8.4	826	22	AAU01366
22	124.5	8.4	10182	23	ABP38314
23	115.5	7.8	1233	21	ABP18246
24	115	7.7	157	23	AAU75479
25	114.5	7.7	674	22	AAU92775
26	113.5	7.6	550	23	ABU53650
27	113	7.6	435	23	ABU55618
28	111	7.5	461	14	AAU39354
29	111	7.5	461	20	AAU43441
30	110.5	7.4	1231	20	AAU00219
31	110.5	7.4	1231	23	ABP43438
32	110.5	7.4	1265	20	AAU00218
33	110.5	7.4	1265	23	ABP43437
34	109.5	7.4	941	22	ABG25801
35	109.5	7.4	1300	22	AAU33407
36	109.5	7.4	1300	22	AAU35313
37	108.5	7.3	287	18	AAU18559
38	108.5	7.3	918	20	AAU08640
39	108.5	7.3	1487	16	AAU74634
40	108	7.3	387	10	AAU90119
41	108	7.3	1023	22	ABG22883
42	108	7.3	1023	22	AAU79772
43	108	7.3	1023	22	AAU79773
44	108	7.3	1076	23	AAU74355
45	108	7.3	1227	22	AAU65637

ALIGNMENTS

RESULT 1

AAU01365

ID AAY01365 standard; protein; 293 AA.

XX AAY01365;

XX 03-JUN-1999 (first entry)

XX Wild-type Staphylococcal alpha haemolysin (HL) polypeptide.

XX Staphylococcal; alpha-haemolysin; alphaHL; mutant; metal; biosensor;
XX heptameric pore assembly; micronutrient analysis; industrial effluent;
XX organic compound; explosive; macromolecule; bacteria; virus.

XX Staphylococcus aureus.

XX WO905167-A1.

XX 04-FEB-1999.

XX 24-JUL-1998; 98WO-US15354.

XX 25-JUL-1997; 97US-0053737.

XX (UYMA-) UNIV MASSACHUSETTS.

XX Bayley H, Braha O, Gouaux E, Kasianowicz J;

XX WPI; 1999-153311/13.

XX New mutant staphylococcal alpha-haemolysin - comprises a
XX heterologous amino acid that binds to analyte, particularly metal
XX ions

Staphylococcus aur
Staphylococcus aur
Staphylococcus aur
(Pro)leukocidin F.
Staphylococcus aur
Staphylococcus aur
Staphylococcus aur
(Pro)leukocidin S
Staphylococcus aur
Alpha haemolysin (S. epidermidis ope
Staphylococcus epi
Plasmidium falcipa
S. aureus antigeni
Human protein sequ
Lactococcus lactis
Lactococcus lactis
EpiP protein. Sta
S. epidermis readi
Enterococcus faeca
E faecalis EF108 a
Enterococcus faeca
E faecalis EF108 p
Novel human diagno
Enterococcus faeca
Enterococcus faeca
Novel fusion prote
S. aureus ClfB pro
Bacillus circulans
Heat stable sarcos
Novel human diagno
Human protein SEQ
Human protein SEQ
Human cytoskeleton
Novel protein kina

XX Disclosure; Page 12; 51pp; English.

XX The invention relates to a mutant staphylococcal alpha-haemolysin

CC (alphaHL) polypeptide. The mutant alphaHL includes a heterologous amino

CC acid (HAA) that binds an analyte, and wherein the polypeptide assembles

CC into a heteroheptameric pore assembly in presence of many wild-type

CC alphaHL polypeptides. The mutant alphaHL polypeptide has at least two

CC non-consecutive HAA in its stem domain, each of which binds: (i) a metal,

CC or (ii) an organic compound. The HAA occupy two or more of the amino acid

CC positions selected from residues 110-149 of the wild-type alphaHL (the

CC present sequence). The HAA can be selected from the group consisting of

CC (i) Ser, Thr, Met, Trp and Tyr or (ii) Glu, Asp, Cys or especially His.

CC Digital biosensors (comprising heptameric pore assemblies containing the

CC mutant alphaHL) are particularly used to detect and quantify metal ions

CC (specifically zinc, cobalt, nickel and cadmium), e.g. in water (for

CC micronutrient analysis), sediment, air, industrial effluent. Organic

CC compounds that can be detected are specifically explosives, but may also

CC be macromolecules or entire bacteria or viruses. The present sequence

CC represents a wild-type staphylococcal alpha haemolysin polypeptide.

XX Sequence 293 AA;

XX Query Match 100.0%; Score 1485; DB 20; Length 293;

XX Best Local Similarity 93.2%; Pred. No. 1.5e-113;

XX Matches 273; Conservative 20; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADSDINIKTGTDTIGSNTTVKTDLVYDKENGMMHKVYFSFIDDKNNHKKLLVIRTKGT 60

DB 1 ADSDINIKTGTDTIGSNTTVKTDLVYDKENGMMHKVYFSFIDDKNNHKKLLVIRTKGT 60

QY 61 IAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTKYXSXLXYXF 120

DB 61 IAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTKYXSXLXYXF 120

QY 121 XGVYXGDXGXIXGXIXAXVXIXHLXYYXQDPFKTILESPT-DKKVGKVKVFNWVNQNW 180

DB 121 NGVNTGDDTKIGGLIGANVSIGHTLUKYVQDPFKTILESPTDKKVGKVKVFNWVNQNW 180

QY 181 PYDRDSWNPVYGNGLPFMKTRNGSMKAADNFDLPNKAASSLLSGFSPDFATVITMDRKASK 240

DB 181 PYDRDSWNPVYGNGLPFMKTRNGSMKAADNFDLPNKAASSLLSGFSPDFATVITMDRKASK 240

QY 241 QQTNDIVYVRDDYQLHWTSTNNKGTNTKDKWTRSSERYKIDWEKEEMTN 293

DB 241 QQTNDIVYVRDDYQLHWTSTNNKGTNTKDKWTRSSERYKIDWEKEEMTN 293

RESULT 2

ID AAU33797

XX AAU33797 standard; Protein; 330 AA.

XX AC AAU33797;

XX 14-FEB-2002 (first entry)

XX Staphylococcus aureus cellular proliferation protein #73.

XX Antisense; prokaryotic cellular proliferation protein;

XX antibiotic; antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

XX 23-MAY-2000; 2000US-206848P.

XX 26-MAY-2000; 2000US-207727P.

XX 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GV;

PI Yamamoto RT, Xu HH;

XX WPI: 2001-611495/70.

DR N-PSDB; AA551656.

XX New polynucleotides for the identification and development of

XX antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 5293; 51pp; English.

XX The invention relates to antisense inhibitors of genes essential to

XX prokaryotic cellular proliferation, their use in identifying the

XX genes, their use in the discovery of novel antibiotics, the essential

XX genes themselves and the encoded proteins. The prokaryotes used are

XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

XX invention is also useful for the identification of potential new targets

XX for antibiotic development. The antisense nucleic acids can also be used

XX to identify proteins used in proliferation, to express these proteins,

XX and to obtain antibodies capable of binding to the expressed proteins.

XX The proteins can be used to screen compounds in rational drug discovery

XX programmes. The antisense nucleic acid sequence is also useful to screen

XX for homologous nucleic acids which are required for cell proliferation in

XX a wide variety of organisms. The present sequence represents an

XX essential prokaryotic cellular proliferation protein.

XX Note: The sequence data for this patent did not form part

XX of the printed specification, but was obtained in electronic

XX format directly from WIPO at

XX ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 330 AA;

XX Query Match 28.6%; Score 424.5; DB 22; Length 330;

XX Best Local Similarity 30.8%; Pred. No. 5e-26;

XX Matches 91; Conservative 80; Mismatches 107; Indels 17; Gaps 8;

QY 7 IKGTGTD-----IGSNTTVKTDLVYDKENGMMHKVYFSFIDDKNNHKKLLVIRTKGTI 61

DB 30 IKTEITPVVEKAIDSETKMYTRTATSDTEKNISQSLQFNFLNEPNYDKETLFIKAKGTI 89

QY 62 AGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTKYXSXLXYXF 120

DB 90 SSGLEILYPNGYWSSTLRWPGSYSVSIQNVDDNTNTKVTDFAPKNQDTRDVKYTYGKT 149

QY 121 XGVYXGDXGXIXGXIXAXVXIXHLXYYXQDPFKTILESPT-DKKVGKVKVFNWVNQNW 179

DB 150 GGNFT-INQNGLSGNITQEYNYSETISYQQPSYRTLIDKPTSNKAVAMKVEANIKNGY 208

QY 180 GPYDRDS---WNPVYGNGLPFMKTRNGSMKAADNFDLPNKAASSLLSGFSPDFATVITMDR 236

DB 209 -DHTRDATDGTGNTKVGSEIFLTRNGNLWARDNFTPKKMPVTVSEGNPEFLTVMHSHDK 267

QY 237 KASKOQTNDIVYVRDDYQLHWTSTN---WKGNTKDKWTRSSERYKIDWE 287

DB 268 K-KGKHSLFVVRKVRMDDFEIRWNYRNWGYWSGKNVNNKKEESLAALIYEIDWK 321

RESULT 3

ID AAU36528

XX AAU36528 standard; Protein; 333 AA.

XX AC AAU36528;

XX 14-FEB-2002 (first entry)

XX Staphylococcus aureus cellular proliferation protein #698.

CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 327 AA;

Query Match 27.2%; Score 403.5; DB 22; Length 327;
Best Local Similarity 28.8%; Pred. No. 2.7e-24;
Matches 78; Conservative 75; Mismatches 113; Indels 5; Gaps 4;

QY 18 TTVKGTDLVTYDKENGHMKVYFSFIDDKHNKLLVIRTKGTIAGQYRVYSEEGANKSG 77
| | | | | : : : : : | | | | | : : : : : | | | : : : |
D 43 TLYKTTATSDNDKLN-ISQILTFENFKDSYDKDVLVKAAGNINSYKPKPNKDYNYSQ 101
| | | | | : : : : : | | | | | : : : : : | | | : : : |

QY 78 LAWPSAFKVLQLPDNEVAQISDYPRNSIDTKYXSXLYXFXGXVXGXGXIXGXIX 137
| | | | | : : : : : | | | | | : : : : : | | | : : : |
D 102 FYWGGKYNVSSESNDVAVVDPYAPKNQNEEFQVQQTGLGYSYGGDI--NISNLSGGLN 159
| | | | | : : : : : | | | | | : : : : : | | | : : : |

QY 138 AXVXIXHXLXYPDFKTIPTD-KKVGKVIFFNNVNMVNMVNMVNMVNMVNMVNMV 196
| | | | | : : : : : | | | | | : : : : : | | | : : : |
D 160 GSKSFSETINKQESYRTTIDRKNHKSIGWVEAHKIMNMGWGPYGRDSDPTYGNELF 219
| | | | | : : : : : | | | | | : : : : : | | | : : : |

QY 197 MKTRNGSKAADNPLDPNKASSLLSGFSPDFATVITMDRKASKQQTNDIVYERVRDDY 256
| | | | | : : : : : | | | | | : : : : : | | | : : : |
D 220 LGGROSSNAGQFLPHTHOMPLLAGNENFEFISVLSHKQNDYK-KSKIKVTYQREMDRY 278
| | | | | : : : : : | | | | | : : : : : | | | : : : |

QY 257 QLHWTSTNWGTNTKDKWTDSSERYKIDWE 287
| | | | | : : : : : | | | | | : : : : : | | | : : : |
D 279 TNQWNLHWYGNVKNQNTVTFTSTVEVDWQ 309
| | | | | : : : : : | | | | | : : : : : | | | : : : |

RESULT 10
AAU37136
ID AAU37136 standard; Protein; 327 AA.
XX
AC AAU37136;
XX
DT 14-FEB-2002 (first entry)
XX
XX Staphylococcus aureus cellular proliferation protein #1306.
DE Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
XX Staphylococcus aureus.
OS
XX WO200170955-A2.
PN 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
PF
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
PI Yamamoto RT, Xu HH;
PI WPI; 2001-611495/70.
DR N-PSDB; AAS4995.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -

XX
PS
XX Example 3; Seq ID No 12729; 511pp; English.
CC
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 327 AA;

Query Match 27.2%; Score 403.5; DB 22; Length 327;
Best Local Similarity 28.8%; Pred. No. 2.7e-24;
Matches 78; Conservative 75; Mismatches 113; Indels 5; Gaps 4;

QY 18 TTVKGTDLVTYDKENGHMKVYFSFIDDKHNKLLVIRTKGTIAGQYRVYSEEGANKSG 77
| | | | | : : : : : | | | | | : : : : : | | | : : : |
D 43 TLYKTTATSDNDKLN-ISQILTFENFKDSYDKDVLVKAAGNINSYKPKPNKDYNYSQ 101
| | | | | : : : : : | | | | | : : : : : | | | : : : |

QY 78 LAWPSAFKVLQLPDNEVAQISDYPRNSIDTKYXSXLYXFXGXVXGXGXIXGXIX 137
| | | | | : : : : : | | | | | : : : : : | | | : : : |
D 102 FYWGGKYNVSSESNDVAVVDPYAPKNQNEEFQVQQTGLGYSYGGDI--NISNLSGGLN 159
| | | | | : : : : : | | | | | : : : : : | | | : : : |

QY 138 AXVXIXHXLXYPDFKTIPTD-KKVGKVIFFNNVNMVNMVNMVNMVNMVNMVNMV 196
| | | | | : : : : : | | | | | : : : : : | | | : : : |
D 160 GSKSFSETINKQESYRTTIDRKNHKSIGWVEAHKIMNMGWGPYGRDSDPTYGNELF 219
| | | | | : : : : : | | | | | : : : : : | | | : : : |

QY 197 MKTRNGSKAADNPLDPNKASSLLSGFSPDFATVITMDRKASKQQTNDIVYERVRDDY 256
| | | | | : : : : : | | | | | : : : : : | | | : : : |
D 220 LGGROSSNAGQFLPHTHOMPLLAGNENFEFISVLSHKQNDYK-KSKIKVTYQREMDRY 278
| | | | | : : : : : | | | | | : : : : : | | | : : : |

QY 257 QLHWTSTNWGTNTKDKWTDSSERYKIDWE 287
| | | | | : : : : : | | | | | : : : : : | | | : : : |
D 279 TNQWNLHWYGNVKNQNTVTFTSTVEVDWQ 309
| | | | | : : : : : | | | | | : : : : : | | | : : : |

RESULT 11
AAU37564
ID AAU37564 standard; Protein; 327 AA.
XX
AC AAU37564;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #1734.
XX Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
XX Staphylococcus aureus.
OS
XX WO200170955-A2.
PN 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.

```

PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI: 2001-611495/70.
XX N-PSDB; AAS55423.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
PT
XX Example 3; Seq ID No 13157; 51lpp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 327 AA;
XX
Query Match 27.2%; Score 403.5; DB 22; Length 327;
Best Local Similarity 28.8%; Pred. No. 2.7e-24;
Matches 78; Conservative 75; Mismatches 113; Indels 5; Gaps 4;
QY 18 TTVKTGDLVTYDKENGMMHKVFYSFIDDKNNHKKLLVIRTKGTIAGQYRVYSEEGANKSG 77
DB 43 TLYKTTATSDNKLN-ISQILFFNIKSDYDKDTLVKLAAGNINSYKPKPKDYNYSQ 101
QY 78 LAWPSAFKVLQLPDNEVAQISDYYPNRSIDTKYXSXLXYFXGXGXDXGXIX 137
DB 102 FYWGKYNVSSESNDVAVVYAPKQNEEFQQOILGYSYGDI--NISNGLSGGLN 159
QY 138 AXVXIXHXLXYXQDPFKTILESPTD-KVGVKVIFFNMVNQWNGPYDRDSMNPVYGNOLF 196
DB 160 GSKSFSETINYKQESYRTTIDRKNHKSIGWGVGAHKNMNGWGPYGRDSDPTVGNELF 219
QY 197 MKTRNGSKAADNFDLPNKASLLSSGSPDPATVITWDRKASKOQTNDIVYIYERVDY 256
DB 220 LGGROSSNAGCNFLPTHOMPLLAGNENPEFISVLSHKQNDTK-KSKIKVTYQREMDRY 278
QY 257 QJHWTSTNKGKNTKDKWTRDSSERYKIDWE 287
DB 279 TNQWNRHLHWGNNYKNQNTVFTSYEVDWQ 309
RESULT 12
AAU34395
ID AAU34395 standard; Protein; 323 AA.
XX
XX AAU34395;
XX
XX 14-FEB-2002 (first entry)

```

```

XX
DE Staphylococcus aureus cellular proliferation protein #671.
XX
XX Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
XX Staphylococcus aureus.
OS
XX WO200170955-A2.
PN
XX 27-SEP-2001.
PD
XX
XX 21-MAR-2001; 2001WO-US09180.
PF
XX
XX 21-MAR-2000; 2000US-191078P.
PR
XX 23-MAY-2000; 2000US-208848P.
PR
XX 26-MAY-2000; 2000US-207727P.
PR
XX 23-OCT-2000; 2000US-242578P.
PR
XX 27-NOV-2000; 2000US-253625P.
PR
XX 22-DEC-2000; 2000US-257931P.
PR
XX 16-FEB-2001; 2001US-269308P.
XX (ELIT-) ELITRA PHARM INC.
PA
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI: 2001-611495/70.
XX N-PSDB; AAS52254.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
PT
XX Example 3; Seq ID No 5891; 51lpp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 323 AA;
XX
Query Match 26.7%; Score 397; DB 22; Length 323;
Best Local Similarity 28.5%; Pred. No. 8.9e-24;
Matches 81; Conservative 83; Mismatches 108; Indels 12; Gaps 7;
QY 13 DIGSNNTVKTGDLVTYDKENGMMHKVFYSFIDDKNNHKKLLVIRTKGTIAGQYRVYSEEG 72
DB 36 NLDGDKMYTATTSDSQKNITSLQNFITEPNYDKETVFIKAKGHTIGSLRILDPNG 95
QY 73 ANKSLGAWPSAFKVLQ-LPDNEVAQISDYYPNRSIDTKYXSXLXYFXGXGXDXGX 131
DB 96 YWNSTLRWPGSYSVSIQVNDNNNTNVTDFAPKNQDESREVKYTYGKGTGD-PSINRG 154
QY 132 IXGXIXAXVXIXHXLXYXQDPFKTILESPTDK-KVGVKVIFFNMVNQWNGPYDRDSW 187
DB 155 LTGNITKESNYSETISYQOQPSYRLLDQSTSHKGVGWKVEAHLINNMGHDTHTQLTNDSD 214

```

QY 188 NPYVGNQFMKTRNGSKAAADNFDLPNKASSLLSSGSPDFATVITMDRKASKQQTNDIV 247
 Db 215 NRT-KSEIFSLTRNGNLWAKDNFTPKDKMPVTYSEGFNPEFLAVMSHDKK-DRGKSOFFV 272
 QY 248 IYERVRDDYQLHWTSTN---WKGNTKDKWTDSSERYKIDWE 287
 Db 273 HYKRMDEFKIDWNHRHGFYSGENHVDKKEKLSALYEVDWK 316

RESULT 13
 AAU37509
 ID AAU37509 standard; Protein; 338 AA.
 XX
 AC AAU37509;
 DT 14-FEB-2002 (first entry)
 XX
 XX Staphylococcus aureus cellular proliferation protein #1679.
 DE Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 KW Staphylococcus aureus.
 OS
 XX WO200170955-A2.
 XX 27-SEP-2001.
 XX 21-MAR-2001; 2001WO-US09180.
 XX 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX (ELIT-) ELITRA PHARM INC.
 XX
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 XX WPI; 2001-611495/70.
 DR N-PSDB; AAS55368.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 XX Example 3; Seq ID No 13102; 511pp; English.
 XX
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 338 AA;
 SQ

Query Match 26.7%; Score 397; DB 22; Length 338;
 Best Local Similarity 28.5%; Pred. No. 9.8e-24;
 Matches 81; Conservative 83; Mismatches 108; Indels 12; Gaps 7;
 QY 13 DIGSNTTVTGDLYTDKENGHKKVPYSPFIDDKNINKLLVIRGTGTAQYRVYSEEG 72
 Db 42 NLDGDTKMTVTRTATSDSQNKITQSLQFNPLTEPNYDKETVFKAKGTIGSLRILDPNG 101
 QY 73 ANKSGLAWPSAFKVLQ-LPDNEVAOISDYPRNSIDTKYXSXLXYFXGVXGXDXGX 131
 Db 102 YWNSTLRWPGSYSVSIQNDNNNTNFTDFAPKNQDSREVKYTYGKTGGD-FSINRGG 160
 QY 132 IXGXIXAXVXIXHXLXYXQPDFKTIILESPDKK-VGKVV---IFNNMNVNMGYPYDRDSW 187
 Db 161 LTGNITKESNYSETISYQPSYRTLLDOSTSHKGVGKVEAHLINNMGHDTTQLTNDSD 220
 QY 188 NPYVGNQFMKTRNGSKAAADNFDLPNKASSLLSSGSPDFATVITMDRKASKQQTNDIV 247
 Db 221 NRT-KSEIFSLTRNGNLWAKDNFTPKDKMPVTYSEGFNPEFLAVMSHDKK-DRGKSOFFV 278
 QY 248 IYERVRDDYQLHWTSTN---WKGNTKDKWTDSSERYKIDWE 287
 Db 279 HYKRMDEFKIDWNHRHGFYSGENHVDKKEKLSALYEVDWK 322

RESULT 14
 AAR35084
 ID AAR35084 standard; Protein; 323 AA.
 XX
 AC AAR35084;
 DT 09-AUG-1993 (first entry)
 DE (Pro)leukocidin F.
 XX
 KW Detection; methicillin-resistant Staphylococcus aureus;
 KW leukocidin; proleukocidin; cancer; antibody.
 XX
 OS Methicillin-resistant Staphylococcus strain 4 (RIMD-3109025).
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..25 /label= sig_peptide
 FT Protein 26..323 /label= mat_protein
 FT
 XX WO9307487-A.
 XX
 PD 15-APR-1993.
 XX
 PF 09-OCT-1992; 92WO-JP01317.
 XX
 PR 09-OCT-1991; 91JP-0290742.
 PR 23-APR-1992; 92JP-0129956.
 XX
 PA (IATR) IATRON LAB INC.
 XX
 PI Izaki K, Kamio Y, Kita H;
 XX
 DR WPI; 1993-134623/16.
 DR N-PSDB; AAQ40216.
 XX
 PT Early detection of methicillin-resistant Staphylococcus aureus -
 PT using antibodies to leukocidin and proleukocidin, useful for
 PT cancer patients, elderly etc.
 XX
 PS Disclosure; Fig 13; 67pp; Japanese.
 XX
 CC DNA and corresp. amino acid sequences of (pro)leukocidin S and F
 CC are claimed. A novel method for the detection of
 CC methicillin-resistant Staphylococcus aureus (MRSA) in biological
 CC samples comprises reaction with an antibody which recognises
 CC (pro)leukocidin S or F.

[illegible]

genes themselves and the encoded proteins. The prokaryotes used are *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*, *Clostridomonas aeruginosa* and *Enterococcus faecalis*. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/publ/pct_sequences](http://wipo.int/pub/publ/pct_sequences).

Query Match	17.0%	Score 252:	DB 22;	Length 321;
Best Local Similarity	24.9%;	Pred. No. 7.9e-12;		
Matches	71;	Conservative	71;	Mismatches 119; Indels 24; Gaps 12;
Qy	13	DIGSNTTV--KTGDLVTDYDKENGHHKVFVSFFDIDKNHKKLLVIRTKGTAGQYRVYSE	70	
Db	47	DICOGAEIIRKTODITS--KRLATQNIQDFVKDKKYNDALVVRMKGFISSR-TTYS	103	
Qy	71	--EGANKSGLAWSAFKVLQLDPNEVAQISDYIPRNISDTKYYXSLXYTFGXVGXD	128	
Db	104	LKKPYIKRMWPFQYNISLTKDSDNY-DUINLPKNKIDSADVSKGLYNGIGNFSAP	162	
Qy	129	XGIXGXIXAAXVIXHXLYXPQDFKTLIBLESPDKKVGVKVIENNMYNQNWGYDRDSWN	188	
Db	163	--SIGG--SGSFNVSKTISYNQNYITEVESQNSKGKVGKVKANSFVTPN-----GV	211	
Qy	189	PVTGNQLFMKTRNGSKAAANFLDPNKAESLLSGSGSPDFATVITMDR-KASQQQNIDV	247	
Db	212	SAYDQYLQAQDPTGP-AARDYFPVDNPQLPPLIOGSNGPSFITLSHERGGDKSE--FEI	268	
Qy	248	IYERVDDYQLHWTTSTNMWGTNTKDWTDRS-SERYKIDWEKEEM	291	
Db	269	TYGRNMDATYAVYTRHRLVADRKHDFAFKNNRTVKVEVNWKTHFV	313	

Search completed: March 4, 2003, 10:42:22
Job time : 35 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:41:14 ; Search time 20.5 seconds
(without alignments)
420.532 Million cell updates/sec

Title: CLAIM6
Perfect score: 1485
Sequence: 1 ADSIDINIKTGTIGSNTV.....WDRSSRYKIDWEKEMTN 293

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2.6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2.6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2.6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2.6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2.6/ptodata/1/1aa/PTUS_COMB.pep.*
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	309	20.8	336	2	US-08-666-405-28
2	124.5	8.4	10182	4	US-09-134-001C-3159
3	111	7.5	461	2	US-08-392-625-24
4	111	7.5	461	2	US-08-466-961A-24
5	111	7.5	461	2	US-08-645-193B-26
6	110.5	7.4	1231	4	US-09-071-035-420
7	110.5	7.4	1265	4	US-09-071-035-418
8	108	7.3	1751	4	US-09-136-574A-44
9	106	7.1	387	4	US-09-457-302-1
10	105.5	7.1	432	1	US-08-476-008-61
11	105.5	7.1	432	1	US-08-306-063-61
12	105.5	7.1	432	1	US-08-833-485-61
13	105.5	7.1	432	4	US-09-137-440-61
14	105	7.1	1584	4	US-09-457-040B-27
15	104.5	7.0	279	4	US-09-134-001C-4878
16	103.5	7.0	440	1	US-08-061-062A-6
17	103.5	7.0	440	1	US-08-061-062A-8
18	103.5	7.0	440	3	US-08-536-150-6
19	103.5	7.0	440	3	US-08-536-150-8
20	101.5	6.8	345	4	US-09-073-898-140
21	101.5	6.8	771	3	US-08-434-000A-8
22	101.5	6.8	771	4	US-09-312-157-8
23	100.5	6.8	686	4	US-09-368-169-8
24	100	6.7	513	1	US-08-200-232-4
25	100	6.7	513	5	PCT-US95-02219-4
26	100	6.7	513	5	PCT-US95-02219A-4
27	99	6.7	537	4	US-09-110-959A-4

28	98	6.6	447	4	US-09-627-376-10	Sequence 10, Appl
29	98	6.6	464	4	US-09-426-072-2	Sequence 2, Appl
30	98	6.6	721	4	US-09-390-234-20	Sequence 20, Appl
31	98	6.6	971	2	US-08-724-354D-22	Sequence 22, Appl
32	98	6.6	971	3	US-09-270-984A-22	Sequence 22, Appl
33	98	6.6	971	3	US-09-177-431-8	Sequence 8, Appl
34	97.5	6.6	1007	4	US-08-961-083-216	Sequence 216, Appl
35	97	6.5	869	2	US-08-483-101-15	Sequence 15, Appl
36	96.5	6.5	4536	4	US-09-180-422B-27	Sequence 27, Appl
37	96	6.5	480	1	US-07-752-428E-2	Sequence 2, Appl
38	96	6.5	480	1	US-07-752-428C-2	Sequence 2, Appl
39	96	6.5	480	1	US-07-752-428C-4	Sequence 4, Appl
40	96	6.5	566	1	US-08-415-823-4	Sequence 4, Appl
41	96	6.5	566	2	US-09-086-662-4	Sequence 4, Appl
42	95	6.4	2314	4	US-09-268-347-49	Sequence 49, Appl
43	94.5	6.4	567	3	US-08-646-273-19	Sequence 19, Appl
44	94.5	6.4	708	3	US-08-646-273-23	Sequence 23, Appl
45	94.5	6.4	754	3	US-08-646-273-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-08-666-405-28
; Sequence 28, Application US/08666405
; Patent No. 5874220
; GENERAL INFORMATION:
; APPLICANT: FACH, Patrick; GUILLLOU,
; APPLICANT: Jean-Pierre; POPOFF, Michel
; TITLE OF INVENTION: PRIMERS FOR THE
; TITLE OF INVENTION: AMPLIFICATION OF GENES CODING FOR THE
; TITLE OF INVENTION: ENTEROTOXIN AND THE LEICITHINASE OF CLOSTRIDIUM
; TITLE OF INVENTION: PEPFRINGENS AND THEIR APPLICATION TO THE
; TITLE OF INVENTION: DETECTION AND NUMERATION OF THESE BACTERIAE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN & MUSERLIAN
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/666,405
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/04292
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/172,026
; FILING DATE: 22-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: MUSERLIAN, CHARLES A
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 102,164
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:

[illegible]


```

Query Match      7.4%; Score 110.5; DB 4; Length 1231;
Best Local Similarity 20.0%; Pred. No. 10;
Matches 55; Conservative 51; Mismatches 108; Indels 61; Gaps 11;

QY 40 YSFIDKHNHKKLLVIRTKGTIAG-----QYRVYSEEGANKSGLAWPSAFK- 85
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 395 YNLTLNENNKKTWSSNLTGDTATLKNLSVPVNYFQYRVNTFYGASDITFTLPKRYKS 454
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 86 VQLQLPDPNEVAQISDYIPRNSID--TKXXYSXLXLYFXGXVXGMDGXIXGXIX 143
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

[illegible]


```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,008
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10660)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-476-008-61

```

```

Query Match      7.1%; Score 105.5; DB 1; Length 432;
Best Local Similarity 20.1%; Pred. No. 3;
Matches 55; Conservative 46; Mismatches 89; Indels 83; Gaps 11;

QY 46 KNHNKLLVTRTGTAQYRVYSEEGANKSGLAWPSAFKVLQLPNEVNAQISDYVPRN 105
   :| | | | | : : : : :| | | | | : : : : :| | | | | :
Db 147 ENEGYPPLAIRNRKGIKGGVKI--DGSISQSFUTALLMSAPLAENDTEIBIIGELVSKP 203
   :| | | | | : : : : :| | | | | : : : : :| | | | | :
QY 106 SID-----TKYXSXLAYXFXGXVKGDXGXIXGXIXAV 140
   || | | | | : : : : :| | | | | : : : : :| | | | | :
Db 204 YIDITLMMRDFGVKVENHHYQFQVKGNSQISPNKYLVEGDASSASYFLAAGAIGKV 263
   :| | | | | : : : : :| | | | | : : : : :| | | | | :
QY 141 XIXH-XLIXYQPD--FKTIILESPTDKKGKVKFNFMVNWQNGPYDRDSWNPVYGNOLFM 197
   :| | | | | : : : : :| | | | | : : : : :| | | | | :
Db 264 KVTGIGKNSIQGRLEFADVLE----KMGAKI-----TWG-----EDFI 297
   :| | | | | : : : : :| | | | | : : : : :| | | | | :
QY 198 KTRNGSKMAADNFDLPNPKASSLLSGFSPFATVITMDRKASKQOOTNIDVIYERVDDYQ 257
   :| | | | | : : : : :| | | | | : : : : :| | | | | :
Db 298 QAEHAELNGID--MDMNH-----PDAAMTIAATLFSNGETVIRNIY----- 338
   :| | | | | : : : : :| | | | | : : : : :| | | | | :
QY 258 LHWTSNWKNTGNTKDKWTDRSSERYKIDWEKEE 290
   || | | | | : : : : :| | | | | : : : : :| | | | | :
Db 339 -----NWRVKET--DRLTAMATELRKKGAVEE 364
   :| | | | | : : : : :| | | | | : : : : :| | | | | :

```

```

RESULT 11
US-08-306-063-61
; Sequence 61, Application US/08306063
; Patent No. 5633435
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis

```

```

; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,063
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(106660)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-306-063-61

Query Match 7.1%; Score 105.5; DB 1; Length 432;
Best Local Similarity 20.1%; Pred. No. 3;
Matches 55; Conservative 46; Mismatches 89; Indels 83; Gaps 11;

QY 46 KNKNKLLVIRTGTTAGQYRVYSEEGANKSGLAWPSAFKVOQLQDPNEVAQISDYYPN 105
   :| || || :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 147 ENEGYPPLARNKGIKGVKVI---DGSISQQLTALLMSAPLAENDTEIEIGELVSKP 203
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 106 SID-----TKXXYSXILXYFXGVXGXDXGXIXGXIXAXV 140
   || :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 204 YIDITLAMRDFGVKVENHHYFQVKVGNQSYISPNKYLVEGDASSAYFLAAGAIGKV 263
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 141 XIXH-XLXYQPP--FKTILESTDKKVGKVIFFNNVNNQNWGPYDRDSWNPYGNQLEM 197
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 264 KVTGIGKNSIQGRLEADVLE---KMGAKI-----TWG-----EDFI 297
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 198 KTRNGSKAADNFDLPNKASLLSGFSFPDFATVITMDRKASQQTNIDVIYERVRDDYQ 257
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 298 QAEHAELNGID--MDMNH-----PDAAMTIATLFSNGEIVIRNY----- 338
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 258 LHWTSFNWKTNTKDKWTDRSSERYKIDWKEE 290
   || :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 339 -----NWRVKET-DRLTAMATELRKVGAEVEE 364

RESULT 12
US-08-833-485-61
; Sequence 61, Application US/08833485
; Patent No. 5804425
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:42:29 ; Search time 10.5 seconds
(without alignments)
1176.748 Million cell updates/sec

Title: CLAIM6

Perfect score: 1485

Sequence: 1 ADSIDINIKTGTDTGSNTTV.....WDRSSRYKIDWEKEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 188354 seqs, 42170167 residues

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	424.5	28.6	330	10	US-09-815-242-5293
2	424.5	28.6	333	10	US-09-815-242-12121
3	404	27.2	325	10	US-09-815-242-5783
4	404	27.2	325	10	US-09-815-242-12693
5	404	27.2	325	10	US-09-815-242-12779
6	404	27.2	325	10	US-09-815-242-13070
7	403.5	27.2	327	10	US-09-815-242-5901
8	403.5	27.2	327	10	US-09-815-242-13729
9	403.5	27.2	327	10	US-09-815-242-13157
10	397	26.7	323	10	US-09-815-242-5891
11	397	26.7	338	10	US-09-815-242-13102
12	252	17.0	321	10	US-09-815-242-5771
13	252	17.0	321	10	US-09-815-242-12694
14	252	17.0	321	10	US-09-815-242-12721
15	234	15.8	312	10	US-09-815-242-5267
16	109.5	7.4	1300	10	US-09-815-242-4303
17	109.5	7.4	1300	10	US-09-815-242-10906
18	109	7.3	861	9	US-09-820-843A-109
19	106.5	7.2	771	10	US-09-818-247-4

20	106	7.1	1349	10	US-09-815-242-5898	Sequence 5898, Ap
21	106	7.1	1349	10	US-09-815-242-13137	Sequence 13137, A
22	105.5	7.1	309	10	US-09-737-178-121	Sequence 121, App
23	105.5	7.1	432	9	US-09-464-099A-61	Sequence 61, Appl
24	105.5	7.1	432	10	US-09-861-696-61	Sequence 61, Appl
25	105	7.1	978	10	US-09-815-242-5456	Sequence 5456, Ap
26	105	7.1	1001	10	US-09-815-242-12686	Sequence 12686, A
27	105	7.1	1352	9	US-09-784-554B-2	Sequence 2, Appl
28	104	7.0	791	9	US-10-055-364-41	Sequence 41, Appl
29	101.5	6.8	345	10	US-09-850-351A-140	Sequence 140, App
30	101.5	6.8	771	9	US-09-982-107-8	Sequence 8, Appl
31	100	6.7	309	10	US-09-737-178-136	Sequence 136, App
32	100	6.7	812	9	US-10-078-770-126	Sequence 126, App
33	99	6.7	1167	10	US-09-815-242-11522	Sequence 11522, A
34	98	6.6	447	12	US-10-047-676A-10	Sequence 10, Appl
35	98	6.6	941	9	US-09-992-598-353	Sequence 353, App
36	98	6.6	941	9	US-09-989-293A-353	Sequence 353, App
37	98	6.6	941	9	US-09-989-735-353	Sequence 353, App
38	98	6.6	941	9	US-09-990-444-353	Sequence 353, App
39	98	6.6	941	9	US-09-989-730-353	Sequence 353, App
40	98	6.6	941	9	US-09-990-436-353	Sequence 353, App
41	98	6.6	941	9	US-09-991-181-353	Sequence 353, App
42	98	6.6	941	9	US-09-993-687-353	Sequence 353, App
43	98	6.6	941	9	US-09-989-734-353	Sequence 353, App
44	98	6.6	941	9	US-10-028-072-464	Sequence 464, App
45	98	6.6	941	9	US-09-997-653-353	Sequence 353, App

ALIGNMENTS

RESULT 1

US-09-815-242-5293
; Sequence 5293, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5293
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5293
Query Match 28.6%; Score 424.5; DB 10; Length 330;
Best Local Similarity 30.8%; Pred. No. 4.5e-24;


```

RESULT 8
US-09-815-242-12729
  Sequence 12729, Application US/09815242
  Patent No. US20020061569A1
  GENERAL INFORMATION:
    APPLICANT: Haselbeck, Robert
    APPLICANT: Ohlsen, Karl L.
    APPLICANT: zyskind, Judith W.
    APPLICANT: Wall, Daniel
    APPLICANT: Trawick, John D.
    APPLICANT: Carr, Grant J.
    APPLICANT: Yamamoto, Robert T.
    APPLICANT: Xu, H. Howard
    TITLE OF INVENTION: Identification of Essential Genes in
    TITLE OF INVENTION: Prokaryotes
    FILE REFERENCE: ELITRA.011A
    CURRENT APPLICATION NUMBER: US/09/815,242
    CURRENT FILING DATE: 2001-03-21
    PRIOR APPLICATION NUMBER: 60/191,078
    PRIOR FILING DATE: 2000-03-21
    PRIOR APPLICATION NUMBER: 60/206,848
    PRIOR FILING DATE: 2000-05-23
    PRIOR APPLICATION NUMBER: 60/207,727
    PRIOR FILING DATE: 2000-05-26
    PRIOR APPLICATION NUMBER: 60/242,578

```

```
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12729
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12729

Query Match 27.2%; Score 403.5; DB 10; Length 327;
Best Local Similarity 28.8%; Pred. No. 1.7e-22;
Matches 78; Conservative 75; Mismatches 113; Indels 5; Gaps 4;

QY 18 TTVKGTDLVTDKENGHMKVFYSFIDDKHNHKKLLVIRTKGTIAGQYRVYSEEGANKSG 77
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 43 TLYKTTATSDNKLN-ISQILTENFIKDSYDKDVLVKAAGNLSYKPKPNKDYNSQ 101

QY 78 LWPFAFKVOLQPDNEVAQISDYPRNSIDTKYXSLXLYXFXGXVXGXDXGXIXGXIX 137
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 102 FYWGGKYNVSSESNDVAVNVDYAPKNQNEEFQVQOTLGYSYGGDI--NISNLSGGLN 159

QY 138 AXVXIXHLXVXOPDFKTIPTD-KKVGKVIFFNNVNNQNGPYDRDSWNPVYGQOLF 196
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 160 GSKSFSTINYKQESYFTTIDRKTNHKSGICGVEAHKIMNNGWGPYGRDSYDPTYGNELF 219

QY 197 MKTRNGSMKAADNPLDPNKAASSLSGFSPDFATVITMDRKASKQOQTNIDVIYERVDDY 256
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 220 LGGRSSSNAGQNLPHQMPLLARGNFPEFISVLSHKQNDYK-KSKIKVTYQREMDRY 278

QY 257 QLHWTSTNWKTGNTKDKWTDSSERYKIDWE 287
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 279 TNQWNLHWVGNVGNKYKNQNTVTFSTYEVDWQ 309

RESULT 9
US-09-815-242-13157
; Sequence 13157, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12729
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12729

Query Match 26.7%; Score 397; DB 10; Length 323;
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13157
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13157

Query Match 27.2%; Score 403.5; DB 10; Length 327;
Best Local Similarity 28.8%; Pred. No. 1.7e-22;
Matches 78; Conservative 75; Mismatches 113; Indels 5; Gaps 4;

QY 18 TTVKGTDLVTDKENGHMKVFYSFIDDKHNHKKLLVIRTKGTIAGQYRVYSEEGANKSG 77
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 43 TLYKTTATSDNKLN-ISQILTENFIKDSYDKDVLVKAAGNLSYKPKPNKDYNSQ 101

QY 78 LWPFAFKVOLQPDNEVAQISDYPRNSIDTKYXSLXLYXFXGXVXGXDXGXIXGXIX 137
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 102 FYWGGKYNVSSESNDVAVNVDYAPKNQNEEFQVQOTLGYSYGGDI--NISNLSGGLN 159

QY 138 AXVXIXHLXVXOPDFKTIPTD-KKVGKVIFFNNVNNQNGPYDRDSWNPVYGQOLF 196
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 160 GSKSFSTINYKQESYFTTIDRKTNHKSGICGVEAHKIMNNGWGPYGRDSYDPTYGNELF 219

QY 197 MKTRNGSMKAADNPLDPNKAASSLSGFSPDFATVITMDRKASKQOQTNIDVIYERVDDY 256
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 220 LGGRSSSNAGQNLPHQMPLLARGNFPEFISVLSHKQNDYK-KSKIKVTYQREMDRY 278

QY 257 QLHWTSTNWKTGNTKDKWTDSSERYKIDWE 287
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
Db 279 TNQWNLHWVGNVGNKYKNQNTVTFSTYEVDWQ 309

RESULT 10
US-09-815-242-5891
; Sequence 5891, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5891
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5891

Query Match 26.7%; Score 397; DB 10; Length 323;
```

```
Best Local Similarity 28.5%; Pred. No. 5.1e-22;
Matches 81; Conservative 83; Mismatches 108; Indels 12; Gaps 7;

QY 13 DIGSNTTVKTGDLVTDKENGHHKVFYFIDDKNNHKKLLVIRTKGTIAGQYRVYSEEG 72
Db 36 NLDGDTKMYTRATTSDSQKNITQSLOFNFTPEPNYDKETVFIKAGTIGSGLRILDPNG 95
QY 73 ANKSGLAWSAFKVLQ-LPDNEVAQISDYPRNSIDTKYXSLXYFXGXVXGXDXGX 131
Db 96 YWNSTLRWPGSYSVSIQNVDNNNTNVTDFAPKNQDESREVKYTYGYKGGD-FSINRGG 154
QY 132 IXGXIXAXVIXHXLXYXQDPFKTTILESTDDK-K-VGWKV---IFNNMNVQNGPYDRDSW 187
Db 155 LTGNITKESYSETISYQOQPSYRTLLDQSTSHKGVGWKVEAHLINNMGHDHTRQLTNDSD 214
QY 188 NPVYGOLFMTKTRNGSMKAADNFDLPNKRASSLLSSGFSDFATVITMDRKASKOOTNIDV 247
Db 215 NRT-KSEIFSLTRNGNLWAKONFTPKDKMPVTVSGFNPFLAVMSHDKK-DKGKSOQV 272
QY 248 IYERVDYQLHWTSTN----WKGNTKDKWTDGRSERYKIDWE 287
Db 273 HYKRSMDFFKIDNRHGFNGWYSGENHVDKKEKLSALYEVDWK 316

RESULT 11
US-09-815-242-13102
; Sequence 13102, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELTRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13102
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13102

Query Match 26.7%; Score 397; DB 10; Length 338;
Best Local Similarity 28.5%; Pred. No. 5.5e-22;
Matches 81; Conservative 83; Mismatches 108; Indels 12; Gaps 7;

QY 13 DIGSNTTVKTGDLVTDKENGHHKVFYFIDDKNNHKKLLVIRTKGTIAGQYRVYSEEG 72
Db 42 NLDGDTKMYTRATTSDSQKNITQSLOFNFTPEPNYDKETVFIKAGTIGSGLRILDPNG 101
QY 73 ANKSGLAWSAFKVLQ-LPDNEVAQISDYPRNSIDTKYXSLXYFXGXVXGXDXGX 131

Best Local Similarity 28.5%; Pred. No. 5.1e-22;
Matches 81; Conservative 83; Mismatches 108; Indels 12; Gaps 7;

QY 13 DIGSNTTVKTGDLVTDKENGHHKVFYFIDDKNNHKKLLVIRTKGTIAGQYRVYSE 70
Db 47 DIGOGAEIIRKTQDITS--KRLAITQNIQDFVKKYKNDALVYKQGFSSR-TTYS 103
QY 71 --EGANKSGLAWSAFKVLQ-LPDNEVAQISDYPRNSIDTKYXSLXYFXGXVXGXDX 128
Db 104 LKPYIKRMWIPQYINISLTKDSNV-DLINYLPPKNKIDSADYSQKLGYNIGNFQ 162
QY 129 XGXIXGXIXAXVIXHXLXYXQDPFKTTILESTDDK-K-VGWKV---IFNNMNVQNGPYDRDSW 188
Db 163 --SISGG--SGSFNTSKTISYQKNVYEVESQNSKGVGWKVKANSFVTPN-----GQV 211
QY 189 PVYGNOLFMTKTRNGSMKAADNFDLPNKRASSLLSSGFSDFATVITMDR-KASKOOTNIDV 247
```

```
Db 102 YWNSTLRWPGSYSVSIQNVDNNNTNVTDFAPKNQDESREVKYTYGYKGGD-FSINRGG 160
QY 132 IXGXIXAXVIXHXLXYXQDPFKTTILESTDDK-K-VGWKV---IFNNMNVQNGPYDRDSW 187
Db 161 LTGNITKESYSETISYQOQPSYRTLLDQSTSHKGVGWKVEAHLINNMGHDHTRQLTNDSD 220
QY 188 NPVYGOLFMTKTRNGSMKAADNFDLPNKRASSLLSSGFSDFATVITMDRKASKOOTNIDV 247
Db 221 NRT-KSEIFSLTRNGNLWAKONFTPKDKMPVTVSGFNPFLAVMSHDKK-DKGKSOQV 278
QY 248 IYERVDYQLHWTSTN----WKGNTKDKWTDGRSERYKIDWE 287
Db 279 HYKRSMDFFKIDNRHGFNGWYSGENHVDKKEKLSALYEVDWK 322

RESULT 12
US-09-815-242-5771
; Sequence 5771, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELTRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5771
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5771

Query Match 17.0%; Score 252; DB 10; Length 321;
Best Local Similarity 24.9%; Pred. No. 4.3e-11;
Matches 71; Conservative 71; Mismatches 119; Indels 24; Gaps 12;

QY 13 DIGSNTTV--KTGDLVTDKENGHHKVFYFIDDKNNHKKLLVIRTKGTIAGQYRVYSE 70
Db 47 DIGOGAEIIRKTQDITS--KRLAITQNIQDFVKKYKNDALVYKQGFSSR-TTYS 103
QY 71 --EGANKSGLAWSAFKVLQ-LPDNEVAQISDYPRNSIDTKYXSLXYFXGXVXGXDX 128
Db 104 LKPYIKRMWIPQYINISLTKDSNV-DLINYLPPKNKIDSADYSQKLGYNIGNFQ 162
QY 129 XGXIXGXIXAXVIXHXLXYXQDPFKTTILESTDDK-K-VGWKV---IFNNMNVQNGPYDRDSW 188
Db 163 --SISGG--SGSFNTSKTISYQKNVYEVESQNSKGVGWKVKANSFVTPN-----GQV 211
QY 189 PVYGNOLFMTKTRNGSMKAADNFDLPNKRASSLLSSGFSDFATVITMDR-KASKOOTNIDV 247
```

Db 212 SAYDQYLFADPTGP-AARDYFVPDNLPLIQSGFNPSFITTLSHERGKGDSE--FEI 268
QY 248 IYERVRDDYQLHWTSTNWKGTNTKDKWTDRS-SERYKIDWEKEEM 291
Db 269 TYGRNMDATYAYVTRHRLAVDRKHDFAFKNRNVTVKYEVNWKTHEV 313

RESULT 13

US-09-815-242-12694

; Sequence 12694, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12694

; LENGTH: 321

; TYPE: PR

; ORGANISM: Staphylococcus aureus

US-09-815-242-12694

Query Match

Best Local Similarity 17.0%; Score 252; DB 10; Length 321;

Mismatches 71; Conservative 71; Mismatches 119; Indels 24; Gaps 12;

QY 13 DIGSNTTV--KTGDLVTDYKENGMMHKVFYFDDKNHNKLLVIRTKGTIAGQYRVYSE 70

Db 47 DIGQGAETIKRTQDITS--KRLAITQNIQDFVKDKKYNKDALVVKMQGFISRR--TYSYD 103

QY 71 --EGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKYXSLXVYFXGXVXGXD 128

Db 104 LKYPYIKRMWPFQYNISLTKTDSNV-DLINVLPKNKIDSADVSKLGVNIGNFQSQAP 162

QY 129 XGIXGXIXAXVIXHXLYXQDPFKTILESPTDKKVGKVIENNVMVQNWGYPDRDSWN 188

Db 163 --SIGG--SGSFNYSKTSISYQKNYVTEVESQNSKGVKGVKANSEFVTPN-----GVQ 211

QY 189 PVYGNOLFMTKTRNGSKAADNFDLPNKASSLSGSGSPDFATVITMDR-KASKQQTNDIV 247

Db 212 SAYDQYLFADPTGP-AARDYFVPDNLPLIQSGFNPSFITTLSHERGKGDSE--FEI 268

QY 248 IYERVRDDYQLHWTSTNWKGTNTKDKWTDRS-SERYKIDWEKEEM 291

Db 269 TYGRNMDATYAYVTRHRLAVDRKHDFAFKNRNVTVKYEVNWKTHEV 313

RESULT 14

US-09-815-242-12721

; Sequence 12721, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12721

; LENGTH: 321

; TYPE: PR

; ORGANISM: Staphylococcus aureus

US-09-815-242-12721

Query Match

Best Local Similarity 17.0%; Score 252; DB 10; Length 321;

Mismatches 71; Conservative 71; Mismatches 119; Indels 24; Gaps 12;

QY 13 DIGSNTTV--KTGDLVTDYKENGMMHKVFYFDDKNHNKLLVIRTKGTIAGQYRVYSE 70

Db 47 DIGQGAETIKRTQDITS--KRLAITQNIQDFVKDKKYNKDALVVKMQGFISRR--TYSYD 103

QY 71 --EGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKYXSLXVYFXGXVXGXD 128

Db 104 LKYPYIKRMWPFQYNISLTKTDSNV-DLINVLPKNKIDSADVSKLGVNIGNFQSQAP 162

QY 129 XGIXGXIXAXVIXHXLYXQDPFKTILESPTDKKVGKVIENNVMVQNWGYPDRDSWN 188

Db 163 --SIGG--SGSFNYSKTSISYQKNYVTEVESQNSKGVKGVKANSEFVTPN-----GVQ 211

QY 189 PVYGNOLFMTKTRNGSKAADNFDLPNKASSLSGSGSPDFATVITMDR-KASKQQTNDIV 247

Db 212 SAYDQYLFADPTGP-AARDYFVPDNLPLIQSGFNPSFITTLSHERGKGDSE--FEI 268

QY 248 IYERVRDDYQLHWTSTNWKGTNTKDKWTDRS-SERYKIDWEKEEM 291

Db 269 TYGRNMDATYAYVTRHRLAVDRKHDFAFKNRNVTVKYEVNWKTHEV 313

RESULT 15

US-09-815-242-5267

; Sequence 5267, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

Search completed: March 4, 2003, 10:46:23
Job time : 11.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:40:49 ; search time 27.5 seconds
(without alignments)
1024.269 Million cell updates/sec

Title: CLAIM6

Perfect score: 1485

Sequence: 1 ADSIDINIKTGTIGSNTTV.....WTDRSSRYKIDWEKEEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	1485	100.0	319	2 S69209	alpha-toxin precu
2	1481	99.7	319	2 F89887	Alpha-Hemolysin pr
3	409	27.5	325	2 B42334	leucocidin R, comp
4	404	27.2	300	2 B49238	gamma-hemolysin ga
5	404	27.2	325	2 F90043	gamma-hemolysin co
6	402.5	27.1	327	2 B89968	leukotoxin, LukD I
7	399	26.9	338	2 C89991	hypothetical prote
8	396	26.7	325	2 T00161	leucocidin chain F
9	394.5	26.6	326	2 S68224	synergohymenotropl
10	393	26.5	325	2 S32212	leucocidin chain F
11	390	26.3	323	2 JQ1530	leukocidin chain F
12	390	26.3	323	2 JQ0627	leukocidin chain F
13	383	25.8	322	2 JC5469	Panton-Valentine l
14	309	20.8	336	2 I40826	beta-toxin - Clost
15	257	17.3	308	2 JC5468	leucocidin chain I
16	255.5	17.2	351	2 D89991	hypothetical prote
17	252	17.0	309	2 D90043	gamma-hemolysin ch
18	251	16.9	309	2 JC4282	gamma-hemolysin II
19	245	16.5	321	2 C49271	hlga-like protein
20	244	16.4	286	2 C49238	gamma-hemolysin co
21	244	16.4	315	2 PC4078	hlgc-like protein
22	244	16.4	315	2 E90043	gamma-hemolysin co
23	241.5	16.3	311	2 C89968	leukotoxin LukE [I
24	241	16.2	315	2 A49234	leucocidin R S com
25	238	16.0	315	2 JN0626	leucocidin chain S
26	237	16.0	312	2 T00160	leucocidin chain S
27	233	15.7	312	2 S32211	leucocidin chain S
28	201.5	13.6	310	2 S68225	synergohymenotropl
29	121	8.1	1562	2 S53069	probable membrane

30	115.5	7.8	1233	1 G71612	novel serine/threo
31	115.5	7.8	1711	2 C71625	variant-specific s
32	114	7.7	1386	2 AC1533	surface protein (L
33	113.5	7.6	550	2 G86667	hypothetical prote
34	113.5	7.6	796	2 A90541	hypothetical prote
35	113.5	7.6	1787	2 T20160	hypothetical prote
36	113	7.6	417	2 G64417	hypothetical prote
37	113	7.6	435	2 G86907	D-alanyl-D-alanine
38	113	7.6	588	2 B71802	penicillin-binding
39	112.5	7.6	453	2 H70352	mannose-6-phosphat
40	112	7.5	976	2 A42466	alpha-amylase (EC
41	111.5	7.5	230	2 JE0083	luza chaperon - Ph
42	111.5	7.5	569	2 A45624	trophozoite cystel
43	111.5	7.5	928	2 S54699	hemopexin-heme com
44	111	7.5	461	2 S23420	probable subtilisi
45	110.5	7.4	437	2 T03553	probable maltose-b

ALIGNMENTS

RESULT 1

S69209

alpha-toxin precursor - Staphylococcus aureus

C:Species: Staphylococcus aureus

C>Date: 06-Sep-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-2000

C:Accession: S69209

R:Hedengrahn, G. submitted to the EMBL Data Library, October 1992

A:Reference number: S69209

A:Accession: S69209

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-319 <HED>

A:Cross-references: EMBL:X01645; NID:g46763; PIDN:CAA25801.1; PID:g46765

C:Superfamily: leukocidin

C:Keywords: toxin

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-319/Product: alpha-toxin #status predicted <MAT>

Query Match 100.0%; Score 1485; DB 2; Length 319;
Best Local Similarity 93.2%; Pred. No. 6.8e-89;
Matches 273; Conservative 20; Mismatches 0; Indels 0; Gaps 0;

QY	1	ADSDINIKTGTDTIGSNTTVKTKGDLVTDKENGMMHKVFYSFIDDKNNHKKLLVIRTKGT	60
Db	27	ADSDINIKTGTDTIGSNTTVKTKGDLVTDKENGMMHKVFYSFIDDKNNHKKLLVIRTKGT	86
QY	61	IAGQYRVYSEEGANKSGLAWPSAFKVLQQLPDNEVAQISDYYPNSIDTKXXYXSLXLYXF	120
Db	87	IAGQYRVYSEEGANKSGLAWPSAFKVLQQLPDNEVAQISDYYPNSIDTKXXYXSLXLYXF	146
QY	121	KGXVXGDXGXIXGXIXAXVYXHXIXYXQPDFKTTILESTDKKVGKVIFFNNMVQNWG	180
Db	147	NGVNTGDDTGKIGGLIGANVSIGHTLKYVQPDFKTTILESTDKKVGKVIFFNNMVQNWG	206
QY	181	PYDRSDSNPNVYGNQLPFKPKFNGSKAADNFDLPNKASLLSSGFSDFATVITMDRKASK	240
Db	207	PYDRSDSNPNVYGNQLPFKPKFNGSKAADNFDLPNKASLLSSGFSDFATVITMDRKASK	266
QY	241	QQTNIIDVIYVERVDDQLHWTSTNWKGTNTKDKWTRSSERYKIDWEKEEMTN 293	
Db	267	QQTNIIDVIYVERVDDQLHWTSTNWKGTNTKDKWTRSSERYKIDWEKEEMTN 319	

RESULT 2

F89887

Alpha-Hemolysin precursor [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 01-Mar-2002

C:Accession: F89887

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Sekimizu, Y. ma, A.; Mizutani-Ul, Y.; Kobayashi, N.; Savano, R.; Inoue, R.; Kaito, C.;

A;Cross-references: EMBL:X64389
C;Superfamily: leukocidin
C;Keywords: toxin

	Query Match	27.5%	Score 409;	DB 2;	Length 325;
	Best local Similarity	28.7%;	Pred. No. 7.3e-19;		
	Matches 81;	Conservative 73;	Mismatches 112;	Indels 16;	Gaps
Qy	11	TTDIGSNTTVKTKGDLVYDKENGMHKKVFYSFDDKNHKKLLAVIRTKGPIAGQYVISE 70			
Db	47	TTATADSDKTKISQILT-----FNFIDKSYKDTLVLKAAAGNINSYGERNP 94			
Dv	71	EGANKSGIAMPSPAFKVOIGLPDNEVAOITSDYYPENRNTDTRKYYXSLIYYFPGXVXGXDXG 130			

Db	95	KYDFSKIYGA	KYNSISQ	SND	SNNVVDY	APKNOEEFQVQNTLGYTFGGDI--SISN	152
Qy	131	XIXGXIAXX	KXIXHLYX	XQDFK	TILESP	TD-KYGVKVIENNMYNQNNGPYPDRDSWNP	189
Db	153	GLSGGLNG	AFSETIN	YKQESFRT	TLRSNTY	KNYGVGVEAHKIMNNGWGPYGRDSFHP	212
Qy	190	VYGNLFPMK	TRNGSMK	AAADLP	DNKASSIL	SSGFSDFATVITMDRKASKQNTINDVIY	249
Db	213	TYGNELEF	LAGQSSAY	AGQNFIA	QHOMPL	LSRNFENPEFLSVLS-HRQDGAKKSKITVY	271
Qy	250	ERVDDYQL	HWTSTNNK	GTNTKDK	WTRD	RSERYKIDWEKEEM	291
Db	272	OREMDIYO	IIRWNGFY	WAGANY	KNFKPT	PKSTVEIDWENHKV	313

```

A:Reference number: A49238; MUID:93138806; PMID:84231103
A:Accession: B49238
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-300 <COO>
A:Note: sequence extracted from NCBI backbone (NCBIN:123247, NCBI:123249)
C:Superfamily: leukocidin
C:Keywords: toxin

      Query Match      27.2%; Score 404; DB 2; Length 300;
      Best Local Similarity 28.7%; Pred. No. 1.4e-18;
      Matches 81; Conservative 72; Mismatches 113; Indels 16; Gaps

QY 11 TTDIGSNVTYKGTGLVLYDKENGHMKKVPYSFIDDKHNHKKLLVLIRKGTIAGQYRVSE 70
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 22 TTATADSKFKISQILT-----FNFIDKDSYKDTLVLKATGINSNGFVKPNP 69

QY 71 EGANKSLAWPSAFKVLQLPDNEVAQISDYIPRNSIDTKYXSLXLYFXGNYVXGDXG 130
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 70 NDYDFSKLYGAKYINSISQSDNSYVVDYAPKQNEEFQVQNTLGYTFGGDI--SISN 127

QY 131 XIXGXIXAXXIXKHXLXYQDPDKTILESPD--KKYWKVIENNMVONWGPVDRDSWNP 189
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 128 GLSGGLNGTAFSETINYQESVFTLSRNTYKNVKGWGEAHKINNNGWGPYGRDSFHP 187

QY 190 VYGNQLFMKTRNGSMKAADNFDLPNKASSLLSGFSDFPATVITMDPKASKOOTNIDVYI 249
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 188 TYGNELFLAGRQSSAYAGQNFIAHQMPLLLSRSFNPEFLSVLS-HRQDGAKKSKITVY 246

QY 250 ERYVDYQLHWTSTNKKGTNRDKWTDRSERYKIDWEKEEM 291
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 247 OREMDLYQIRWNGFYWAGANYKNFKTRTFKSTYEIDWENHKV 288

RESULT 5

```

F90043
gamma-hemolysin component B [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: F90043
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.; Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F90043
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <KUR>
A:Cross-references: GB:BA000018; PID:g13702370; PIDN:BAB43511.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: hlgB
C:Superfamily: leukocidin

Query Match 27.2%; Score 404; DB 2; Length 325;
Best Local Similarity 29.4%; Pred. No. 1.6e-18;
Matches 87; Conservative 74; Mismatches 119; Indels 16; Gaps 6;
QY 9 TGTDTI-GSNTTV---KTGDLVTYDKENG-----MHRKKVFSFIDDKNNHKKLLVIR 56
DB 21 SGTANAEGKTPVSVKVKVDKVLTKYTATADSKFKISILFNFKDKSYDKDTLVLK 80
QY 57 TKGTIAGQYRVYSEGANSGSLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKXXYXSL 116
DB 81 ATGNINSFGKPNPDYDFSKLYGAKYVNSISQSDNSVNVVDYAPKQNEEFQVQNTL 140
QY 117 XYFXGAVXGDXGXIXGXIXYXIXYXQDPDKTILESPD-KKVGKVFNNMV 175
DB 141 GYTFGGDI--SISNGLSGLNGTAFSETINYKQESYRTLSRNTYKNYGVGVEAHKIM 198
QY 176 NQWGPYDRDQSWNPVYCNLFMRNGSMKAADNFPDNPKNASSLLSSGFSFDPATVITMD 235
DB 199 NNGWPGYGRDSFHTYFTEUFLAGQSSAYAGQNFIAHQHQPMLLSRNFNFPEFLSVLS-H 257
QY 236 RKASKOOTNIDVIYERVDYQLHWTSTNMKGTNTKDKWDRSSERYKIDWEKEEM 291
DB 258 RDGAKKSKITVTQREMDLYQIRWNGFYWAGANYKFKTKTPKSTVEIDWENHKV 313

RESULT 6
leukotoxin, LukD [Pathogenicity island SapIn3] [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89968
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.; Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B89968
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-327 <KUR>
A:Cross-references: GB:BA000018; PID:g13701612; PIDN:BAB42905.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: lukD
C:Superfamily: leukocidin

Query Match 27.1%; Score 402.5; DB 2; Length 327;
Best Local Similarity 28.8%; Pred. No. 2e-18;
Matches 78; Conservative 75; Mismatches 113; Indels 5; Gaps 4;
QY 18 TTVTGTDLVTYDKENGHKKVFSFIDDKNNHKKLLVIRTKGTIAGQYRVYSEGANSG 77

F90043
gamma-hemolysin component B [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: F90043
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.; Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F90043
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <KUR>
A:Cross-references: GB:BA000018; PID:g13702370; PIDN:BAB43511.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: hlgB
C:Superfamily: leukocidin

Query Match 27.2%; Score 404; DB 2; Length 325;
Best Local Similarity 29.4%; Pred. No. 1.6e-18;
Matches 87; Conservative 74; Mismatches 119; Indels 16; Gaps 6;
QY 9 TGTDTI-GSNTTV---KTGDLVTYDKENG-----MHRKKVFSFIDDKNNHKKLLVIR 56
DB 21 SGTANAEGKTPVSVKVKVDKVLTKYTATADSKFKISILFNFKDKSYDKDTLVLK 80
QY 57 TKGTIAGQYRVYSEGANSGSLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKXXYXSL 116
DB 81 ATGNINSFGKPNPDYDFSKLYGAKYVNSISQSDNSVNVVDYAPKQNEEFQVQNTL 140
QY 117 XYFXGAVXGDXGXIXGXIXYXIXYXQDPDKTILESPD-KKVGKVFNNMV 175
DB 141 GYTFGGDI--SISNGLSGLNGTAFSETINYKQESYRTLSRNTYKNYGVGVEAHKIM 198
QY 176 NQWGPYDRDQSWNPVYCNLFMRNGSMKAADNFPDNPKNASSLLSSGFSFDPATVITMD 235
DB 199 NNGWPGYGRDSFHTYFTEUFLAGQSSAYAGQNFIAHQHQPMLLSRNFNFPEFLSVLS-H 257
QY 236 RKASKOOTNIDVIYERVDYQLHWTSTNMKGTNTKDKWDRSSERYKIDWEKEEM 291
DB 258 RDGAKKSKITVTQREMDLYQIRWNGFYWAGANYKFKTKTPKSTVEIDWENHKV 313

RESULT 6
leukotoxin, LukD [Pathogenicity island SapIn3] [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89968
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.; Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B89968
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-327 <KUR>
A:Cross-references: GB:BA000018; PID:g13701612; PIDN:BAB42905.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: lukD
C:Superfamily: leukocidin

Query Match 27.1%; Score 402.5; DB 2; Length 327;
Best Local Similarity 28.8%; Pred. No. 2e-18;
Matches 78; Conservative 75; Mismatches 113; Indels 5; Gaps 4;
QY 18 TTVTGTDLVTYDKENGHKKVFSFIDDKNNHKKLLVIRTKGTIAGQYRVYSEGANSG 77

Db 43 TLYKTATSDNDKLN-ISQILTENFIKDSYDKDQTLVLKAAGNINSYKKPNPKDYNYSQ 101
QY 78 LWPSPAFKVLQLPDNEVAQISDYPRNSIDTKXXYXSLXYFXGAVXGDXGXIXGXIX 137
Db 102 FYGGKYNVSVSSNESDANVVDYAPKQNEEFQVQNTLGLSYGGDI--NISNGLSGLN 159
QY 138 AXVKIXHXLAYXQDPDKTILESPD-KKVGKVFNNVNMVNMWGPYDRDSWNPVYGOLF 196
Db 160 GSKFSFTINVKESYRTTIDRNTNHSKISGWVEAHKIMNMGWGPYDRSDYDPTYGNELF 219
QY 197 MKTRNGSMKAADNFPDNPKNASSLLSGFSFDPATVITMDRKASKOOTNIDVIYERVDY 256
Db 220 LGGRQSSNAGQNFPLPHQMPLLARGNFPFISVLSHKQNDTK-KSKIKVTYQREMDRY 278
QY 257 QLHWTSTNWKGTNTKDKWDRSSERYKIDWE 287
Db 279 TNQWNLHGWIGNNYKNQNTVTFTSTYEVDWQ 309

RESULT 7
C89991
hypothetical protein SA1812 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89991
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.; Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C89991
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-338 <KUR>
A:Cross-references: GB:BA000018; PID:g13701799; PIDN:BAB43092.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA1812
C:Superfamily: leukocidin

Query Match 26.9%; Score 399; DB 2; Length 338;
Best Local Similarity 28.1%; Pred. No. 3.5e-18;
Matches 81; Conservative 85; Mismatches 110; Indels 12; Gaps 7;
QY 13 DIGSNTTVTKTGDLVTYDKENGHKKVFSFIDDKNNHKKLLVIRTKGTIAGQYRVYSEB 72
Db 42 NLDGDTKMYTRTATTSQSKNITQSLQNFLETPNYDKETVFIKAKGTIGSLRLDPNG 101
QY 73 ANKSGLAWPSAFKVLQ-LPDNEVAQISDYPRNSIDTKXXYXSLXYFXGAVXGDXGX 131
Db 102 YWNSTLWPGYSVSIQNVDDNNNTVDFAPKNQDSREVKYTYGKTGCD-FSINRGG 160
QY 132 IXGXIXAXVIXIXHXLAYXQDPDKTILESPD-KKVGKVFNNVNMWGPYDRDSW 187
Db 161 LTGNITKESYSETISIQQSYRTLLDQSTSHKGVGHKVAHLNNGHDFHRLTNDSD 220
QY 188 NPVYGNOLFMTKNGSMKAADNFPDNPKNASSLLSGFSFDPATVITMDRKASKOOTNIDV 247
Db 221 NRT-KSEIFSLTRNGNLWAKDNFTPKDKMPVTVSEGFNPEFLAVMSHDKK-DKGSQFV 278
QY 248 IYERVDYQLHWTSTN---WKGNTNFKDKWDRSSERYKIDWEKEEM 291
Db 279 HYKRSMDFEKIDWNRHGFNGWGSVENHVDKKEKLSALYEVDWKTHDV 326

RESULT 8
T00161
leukocidin chain F precursor - Staphylococcus aureus phage phi pVL
N; Alternate names: protein 28
C:Species: Staphylococcus aureus phage phi pVL
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000

QY 106 SIDTKYXSLXVFXGXVGDGXGXIXXAXVIXHXLXVXQDPDFKTLSPD-KK 164
Db 130 QNEEFQVQDTGLVSYGGDI--NIINGLTGGLNGSKSFSETINYKQESYRTTIDRKTNHS 187
QY 165 VGMKVFNNMNVONWGPYDRDSDNPNVYGNQLPFKMTGSKMAADNFDLPNKASSLLSSGF 224
Db 188 IGVGVFAHKIMNGWGPYGRDSDSLYGNEFLFGGQSSNANQNQLPHTQPIIARGNF 247
QY 225 SPDFATVITMDRKASKQOQTNIDVIYVRDDYQLHWTSTNWKGTNTKDKWTDKDRSSRYKI 284
Db 248 NPEFISVLSHKQDKVK-KSKIKVYQREMDRYENFNHLHWGYNKKNQKRATHTSIYEL 306
QY 285 DWK 288
Db 307 DWK 310

RESULT 14

I40826
beta-toxin - Clostridium perfringens
C:Species: Clostridium perfringens
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 17-Nov-2000
C:Accession: I40826; I40856; S51649
R:Hunter, J.E.C.; Bronn, J.E.; Oyston, P.C.F.; Jakura, J.; Titball, R.W.
Infect. Immun. 61, 3958-3965, 1993
A:Title: Molecular genetic analysis of the beta-toxin of Clostridium perfringens re-
A:Reference number: I40826; MUID:93366460; PMID:8359918
A:Accession: I40826
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-336 <RES>
A:Cross-references: GB:L13198; NID:9410018; PIDN:AAA23284.1; PID:9410019
R:Steinporstottir, V.; Frithriksdottir, V.; Gunnarsson, E.; Andresson, O.S.
FEMS Microbiol. Lett. 130, 273-278, 1995
A:Title: Expression and purification of Clostridium perfringens beta-toxin glutathione S
A:Reference number: I40856; MUID:93377614; PMID:7649450
A:Accession: I40856
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-170, 'K', 172-336 <RE2>
A:Cross-references: EMBL:X83275; NID:9602993; PIDN:CAA58246.1; PID:9602994
C:Genetics:
A:Gene: cfp
C:Superfamily: leukocidin
C:Keywords: toxin

Query Match 20.8%; Score 309; DB 2; Length 336;
Best Local Similarity 25.3%; Pred. No. 2.5e-12;
Matches 80; Conservative 75; Mismatches 115; Indels 46; Gaps 11;

QY 13 DIGSNTTV---KTGD-----LVTYD-----KENGMMKKVYFSDIDDKNNHKK 51
Db 29 DIGKTTTTRNKTSDGYTIITQNDQIISYQSDSSKNEDEGFTASIDARFIDDKYSSE 88
QY 52 LLVITKGTIAG-----QYRVSEEGANKSGLWPSAFKVO-IQLPDNEVAQISDYYP 103
Db 89 TFLNLTGFMSSKEDVKKYNLH--DVTNSTAINFPVRYISISILNESINENVKIVDSIP 146
QY 104 RNSIDTKYXSLXVFXGXVGDGXGXIXXAXVIXHXLXVXQDPDFKTLSPD-KK 163
Db 147 KNTISQKVTNMGKIGGSIE-IEENPKASIESEYAESSTIEYQDPDFTSIQPDHSTS 205
QY 164 KVGKVFNNMNVONWGPYDRDSDNPNVYGNQLPFKMTGSKMAADNFDLPNKASSLLSSG 223
Db 206 KASWDTKFTETTR---GNVLKSNPNVYGNEMFYGRYTNVPATENIIPDYQMSKLIITGG 262
QY 224 FSPDFATVITMDRKASKQOQTNIDVIYVRDDYQLHWTSTNWKGT-----NTKDKWTD 276
Db 263 LNPNNSVVLTAPN--GTEESIIKVKMERNCYLLNNGANNVGVQVYSLAFDTPN--VD 318
QY 277 RSSERYKIDWEKEMT 292
Db 319 SHIFTKINLWTHKVT 334

RESULT 15

JC5468
leukocidin chain lukM precursor - Staphylococcus aureus
C:Species: Staphylococcus aureus
C>Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 20-Jun-2000
C:Accession: JC5468; PC4338; S68460; S68542
R:Kaneko, J.; Muramoto, K.; Kamio, Y.
Biosci. Biotechnol. Biochem. 61, 541-544, 1997
A:Title: Gene of LukF-PV-like component of Pantan-Valentine leukocidin in Staphylococ-
A:Reference number: JC5468; MUID:97249691; PMID:9095557
A:Accession: JC5468
A:Molecule type: DNA
A:Residues: 1-308 <KAN>
A:Cross-references: DDBJ:D83951; NID:gl230553; PIDN:BAAL2147.1; PID:gl230554
A:Experimental source: strain P83
A:Accession: PC4338
A:Molecule type: protein
A:Residues: 29-75;101-130;263-283 <KA2>
R:Choorit, W.; Kaneko, J.; Muramoto, K.; Kamio, Y.
FEBS Lett. 357, 260-264, 1995
A:Title: Existence of a new protein component with the same function as the LukF comp-
A:Reference number: S68460; MUID:95137101; PMID:7835424
A:Accession: S68460
A:Molecule type: DNA
A:Residues: 1-2, 'N', 4-288, 'L', 290-299, 'S', 301, 'L', 303-304, 'R', 306-308 <CHO>
A:Cross-references: EMBL:D42144; NID:9577648; PIDN:BAA07715.1; PID:9577649
A:Experimental source: strain P83 (ATCC 31890)
A:Accession: S68542
A:Molecule type: protein
A:Residues: 29-75;101-130;263-282 <CHW>
C:Genetics:
A:Gene: lukM
C:Function:
A:Description: causes cytotoxic changes in polymorphonuclear leukocytes in human and
C:Superfamily: leukocidin
C:Keywords: toxin
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-308/Product: leukocidin chain lukM #status experimental <MAT>
Query Match 17.3%; Score 257; DB 2; Length 308;
Best Local Similarity 24.1%; Pred. No. 5.2e-09;
Matches 70; Conservative 74; Mismatches 121; Indels 26; Gaps 12;

QY 9 TCTTDDIGSNTTV--KTGDLVTYDKENGMMKKVYFSDIDDKNNHKKLLVIRTKGTIAG-- 64
Db 30 TNAEDIGDDAEVYIKRTEDVSS--RRWGVTONVQDFVKKYKNDALIRKQGFINSRTT 87
QY 65 YRVYSEEGANKSGLWPSAFKVOQLPDNEVAQISDYYPNRSIDTKYXSLXVFXGXV 124
Db 88 FNDVKQNRANKR--WVWPFQYNIGLTSKQDNTSLI--NYPKKNKIEIVDVQTLGYNIGKGF 145
QY 125 XGXDXGXIXGXIXAXVIXHXLXVXQDPDFKTLSPDTPDKKVGKVFNNMV--NONWGPY 182
Db 146 QSVF--STGG--NGSFNYSKISYKSYVSEVEQSSKTIKGVKANSFVIAGHRWSAY 201
QY 183 DRDSNPNVYGNQLPFKMTGSKMAADNFDLPNKASSLLSSGSPDFAVITMDRKASKQO 242
Db 202 D-----ELLTIRNTRGPNARDYFDDNPLPPLTSGFNPSFIATVSHE--KDSGDT 251
QY 243 TNIDVIYVRD-DYQLHWTSTNWKGTNTKDKWTDKRS--SERIKIDWEKEEM 291
Db 252 SEFETIYGRNMDVIYATYLPKGLYPERKHNEFVNRNRFVYKVEYNWKTYEI 302

Search completed: March 4, 2003, 10:45:09
Job time : 29.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:39:33 ; Search time 17 Seconds
(without alignments)
714.857 Million cell updates/sec

Title: CLAIM6

Perfect score: 1485

Sequence: 1 ADSDNINIKGTGTDIGSNTV.....WDRSSERYKIDWEKEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1485	100.0	319	1 HLA_STAAU	P09616 staphylococ
2	404	27.2	325	1 HLGB_STAAN	Q07226 staphylococ
3	390	26.3	323	1 LUKF_STAAM	P31715 staphylococ
4	252	17.0	309	1 HLGA_STAAM	P31714 staphylococ
5	244	16.4	315	1 HLGC_STAAM	Q07227 staphylococ
6	238	16.0	315	1 LUKS_STAAM	P31716 staphylococ
7	121	8.1	1562	1 YM81_YEAST	Q04781 saccharomyc
8	113.5	7.6	1787	1 CHD3_CAEEL	Q22516 caenorhabdi
9	113	7.6	417	1 Y943_METJA	Q58353 methanococc
10	112	7.5	976	1 AMY_BUTFI	P30269 butyrivibri
11	111.5	7.5	569	1 CYSP_PLAFA	P25805 plasmodium
12	111.5	7.5	928	1 HXA2_HAEIN	P45354 haemophilus
13	111	7.5	461	1 EPIP_STAEP	P30199 staphylococ
14	108.5	7.3	1786	1 YCF1_ARATH	P56785 arabidopsis
15	107.5	7.2	420	1 FAB2_MYCLE	Q69473 mycobacteri
16	106.5	7.2	771	1 PTGR_MOUSE	O70570 mus musculu
17	106.5	7.2	1066	1 HXSA_STRPN	Q54873 streptococc
18	106	7.1	387	1 M5OX_BACSN	P23342 bacillus sp
19	105.5	7.1	380	1 GUG2_YEAST	P47011 saccharomyc
20	105	7.1	464	1 MRJ4_APIME	Q17061 apis mellif
21	105	7.1	859	1 RPA2_METJA	Q58446 methanococc
22	105	7.1	1584	1 KYK1_DICDI	P18160 dictyosteli
23	105	7.1	1655	1 OMPB_RICCN	Q9Kka3 r outer mem
24	104	7.0	1656	1 OMPB_RICUA	Q06653 r outer mem
25	103.5	7.0	440	1 RYGA_ASPAC	Q00001 aspergillus
26	103	6.9	1212	1 YK64_CAEEL	P34343 caenorhabdi
27	103	6.9	1698	1 4L_DROME	Q9v8r9 drosophila
28	102	6.9	784	1 OSTA_ECOLI	P31554 escherichia
29	101.5	6.8	1160	1 TFC3_YEAST	P34111 saccharomyc
30	101	6.8	467	1 RGSB_HUMAN	O94810 homo sapien
31	101	6.8	679	1 RRP1_DROME	P27864 drosophila
32	101	6.8	1007	1 RGA1_YEAST	P39083 saccharomyc
33	101	6.8	1310	1 VAC3_HELPY	Q48253 helicobacte

34	101	6.8	1365	1 GTF5_STRDO	P29336 streptococc
35	100.5	6.8	441	1 YY07_METJA	O60306 methanococc
36	100.5	6.8	556	1 Y281_MYCGE	P47523 mycoplasma
37	100.5	6.8	705	1 PPCE_FLAME	P47028 flavobacter
38	100	6.7	649	1 GUND_CLOTR	Q04954 clostridium
39	99.5	6.7	432	1 AROA_HAEIN	Q03421 haemophilus
40	99	6.7	441	1 YKF9_YEAST	P35728 saccharomyc
41	99	6.7	510	1 INOL_NICPA	Q9ssv4 nicotiana p
42	99	6.7	1018	1 FNBA_STAAM	P14738 staphylococ
43	99	6.7	1592	1 GTF2_STRDO	P27470 streptococc
44	99	6.7	1597	1 GTF1_STRDO	P11001 streptococc
45	99	6.7	1783	1 Y468_MYCGE	Q49460 mycoplasma

ALIGNMENTS

RESULT 1

ID	HLA_STAAU	STANDARD;	PRT;	319 AA.
AC	P09616;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Alpha-hemolysin precursor (Alpha-toxin) (Alpha-HL).			
GN	HLV OR HLA.			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=1280;			
RN	{1}			
RC	SEQUENCE FROM N.A., AND SEQUENCE OF 27-44.			
RP	STRAIN=Wood 46;			
RX	MEDLINE=85053471; PubMed=6500704;			
RA	Gray G.S., Kehoe M.;			
RT	"Primary sequence of the alpha-toxin gene from Staphylococcus aureus			
RT	wood 46.";			
RL	Infect. Immun. 46:615-618(1984).			
RN	{2}			
RP	REVISIONS, SEQUENCE FROM N.A.			
RC	STRAIN=Wood 46;			
RA	Hedengrahn G.;			
RX	Submitted (Oct-1992) to the EMBL/GenBank/DBJ databases.			
RN	{3}			
RP	SEQUENCE OF 27-319 FROM N.A., AND PARTIAL SEQUENCE.			
RX	STRAIN=Wood 46;			
RC	MEDLINE=92268149; PubMed=1587866;			
RX	Walker B., Krishnaswamy M., Zorn L., Kasianowicz J., Bayley H.;			
RA	"Functional expression of the alpha-hemolysin of Staphylococcus			
RT	aureus in intact Escherichia coli and in cell lysates. Deletion of			
RT	five C-terminal amino acids selectively impairs hemolytic			
RT	activity.";			
RL	J. Biol. Chem. 267:10902-10909(1992).			
RN	{4}			
RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).			
RX	STRAIN=Wood 46 / ATCC 10832;			
RC	MEDLINE=97102581; PubMed=8943190;			
RX	Song L., Hobough M.R., Shustak C., Cheley S., Bayley H., Gouaux J.E.;			
RA	"Structure of staphylococcal alpha-hemolysin, a heptameric			
RT	transmembrane pore.";			
RL	Science 274:1859-1866(1996).			
RN	{5}			
RP	MUTAGENESIS.			
RX	MEDLINE=93016135; PubMed=1400487;			
RA	Walker B., Krishnaswamy M., Zorn L., Bayley H.;			
RT	"Assembly of the oligomeric membrane pore formed by staphylococcal			
RT	alpha-hemolysin examined by truncation mutagenesis.";			
RL	J. Biol. Chem. 267:21782-21786(1992).			
RN	{6}			
RP	MUTAGENESIS OF HISTIDINE RESIDUES.			
RX	MEDLINE=94222552; PubMed=8168947;			
RA	Menzies B.E., Kernodle D.S.;			
RT	"Site-directed mutagenesis of the alpha-toxin gene of Staphylococcus			
RT	aureus: role of histidines in toxin activity in vitro and in a murine			

claim6.rsp

Tue Mar 4 12:36:57 2003

FT model.;
 RL Infect. Immun. 62:1843-1847(1994).
 RN [7]
 RP MUTAGENESIS OF HISTIDINE RESIDUES.
 RX MEDLINE=94245329; PubMed=8188346;
 RA Jursch R., Hildebrand A., Hobom G., Tranum-Jensen J., Ward R.,
 RA Kehoe M., Bhakdi S.;
 RA "Histidine residues near the N-terminus of staphylococcal alpha-toxin
 RT as reporters of regions that are critical for oligomerization and
 RT pore formation.";
 RT Infect. Immun. 62:2249-2256(1994).
 RL [8]
 RN MUTAGENESIS.
 RP MEDLINE=96032742; PubMed=7559447;
 RX Walker B., Bayley H.;
 RA "Key residues for membrane binding, oligomerization, and pore forming
 RT activity of staphylococcal alpha-hemolysin identified by cysteine
 RT scanning mutagenesis and targeted chemical modification.";
 RT J. Biol. Chem. 270:23065-23071(1995).
 CC -1- FUNCTION: ALPHA-TOXIN BINDS TO THE MEMBRANE OF EUKARYOTIC CELLS
 CC RESULTING IN THE RELEASE OF LOW-MOLECULAR WEIGHT MOLECULES AND
 CC LEADING TO AN EVENTUAL OSMOTIC LYSIS. HEPTAMER OLIGOMERIZATION
 CC LEADING TO AN EVENTUAL OSMOTIC LYSIS.
 CC AND PORE FORMATION IS REQUIRED FOR LYTIC ACTIVITY.
 CC -1- SUBUNIT: SELF-ASSEMBLES TO FORM FIRST, A NONLYTIC OLIGOMERIC
 CC INTERMEDIATE, AND THEN, A MUSHROOM-SHAPED HOMOPENTAMER STRUCTURE
 CC OF 100 ANGSTROMS IN LENGTH AND UP TO 100 ANGSTROMS IN DIAMETER.
 CC -1- SUBCELLULAR LOCATION: SECRETED AS A MONOMER. AFTER OLIGOMERIZATION
 CC AND PORE FORMATION, THE COMPLEX IS TRANSLOCATED ACROSS THE
 CC BILAYER, PROBABLY VIA THE GLY-RICH DOMAIN OF EACH STRAND. DOMAIN
 CC -1- DOMAIN: THE MUSHROOM-SHAPED HEPTAMER IS COMPOSED OF A EACH
 CC (COMPRISING 7 BETA SANDWICHES AND THE AMINO LATCHES OF EACH
 CC PROTEIN). 7 RIM REGIONS WHOSE PROTRUDING STRANDS MAY INTERACT
 CC WITH THE MEMBRANE BILAYER, AND THE STEM DOMAIN (52 ANGSTROMS IN
 CC LENGTH, 26 ANGSTROMS IN DIAMETER) WHICH FORMS THE TRANSMEMBRANE
 CC PORE.
 CC -1- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL; X01645; CAA25801.1;
 CC DR EMBL; M90536; AAA26598.1;
 CC DR PDB; 7AHL; 14-JAN-98.
 CC DR InterPro: IPR001776; Aerolysin.
 CC DR InterPro: IPR001340; Hemolysn_pore.
 CC DR InterPro: IPR003963; Staph_bich_tkn.
 CC DR Pfam: PF01117; Aerolysin; 1.
 CC DR PRINTS; PR01468; BICOMPNTOXIN.
 CC DR TIGRFS; TIGR01002; hlyII; 1.
 CC DR PROSITE; PS00274; AEROLYSIN; 1.
 CC DR Hemolysis; Toxin; Signal; 3D-structure.
 CC KW SIGNAL
 FT 26 ALPHA-HEMOLYSIN.
 FT 27 GLY-RICH.
 FT 145 CLEAVAGE OF MONOMERS BY PROTEINASE K
 FT 159 CLEAVAGE OF MONOMERS BY PROTEINASE K
 FT SITE (MINOR SITE).
 FT 161 CLEAVAGE OF MONOMERS BY PROTEINASE K
 FT SITE (MAJOR SITE).
 FT 162 CLEAVAGE OF MONOMERS BY PROTEINASE K
 FT SITE (MINOR SITE).
 FT 165 CLEAVAGE OF MONOMERS BY PROTEINASE K
 FT SITE (MAJOR SITE).
 FT 166 H->L: NO OLIGOMERIZATION NOR HEMOLYTIC
 FT MUTAGEN 61 ACTIVITY.
 FT 61 H->R: NO OLIGOMERIZATION NOR HEMOLYTIC
 FT MUTAGEN 61 ACTIVITY.
 FT 74 H->L: 7% OF NORMAL HEMOLYTIC ACTIVITY.
 FT 170 H->L: 16% OF NORMAL HEMOLYTIC ACTIVITY.
 FT MUTAGEN 170

FT MUTAGEN 285 285 H->L: 46% OF NORMAL HEMOLYTIC ACTIVITY.
 SQ SEQUENCE 319 AA; 35904 MW; 6711C415DF7EBF30 CRC64;
 Query Match 100.0%; Score 1485; DB 1; Length 319;
 Best Local Similarity 93.2%; Pred. No. 1.8e-88; Indels 0; Gaps 0;
 Matches 273; Conservative 20; Mismatches 0;
 QY 1 ADSDINIKTGTDTGNTTIVKTDVTVYDKENGMMHKVYFSDIDKNNKKLLVIRTKGT 60
 DB 27 ADSDINIKTGTDTGNTTIVKTDVTVYDKENGMMHKVYFSDIDKNNKKLLVIRTKGT 86
 QY 61 TAGQRYVYSEEGANKSGLWPSAFKVQLQDPNEVAQISDYPRNSIDTKYXSLAYXF 120
 DB 87 TAGQRYVYSEEGANKSGLWPSAFKVQLQDPNEVAQISDYPRNSIDTKYXSLAYXF 146
 QY 121 XGVXGXDXGXIXIXAXXVIXHXLXIXQPDFKILESPDKKYGKVFNNMNNQNG 180
 DB 147 NGNVTDGDTGKIGGLIGANVSIGHTLVYQPDFKILESPDKKYGKVFNNMNNQNG 206
 QY 181 PYDRDSNVPVYGNQLFMTKTRNGSKAANDFLDPNKASSLLSSGFSFDFATVITMDRKASK 240
 DB 207 PYDRDSNVPVYGNQLFMTKTRNGSKAANDFLDPNKASSLLSSGFSFDFATVITMDRKASK 266
 QY 241 QQTNIIDVYERDDYOLHWTSTNNKGTNTKDTDRSSERYKIDWEKEMTN 293
 DB 267 QQTNIIDVYERDDYOLHWTSTNNKGTNTKDTDRSSERYKIDWEKEMTN 319
 RESULT 2
 HLG8 STAA STAA STANDARD; PRT; 325 AA.
 ID HLG8 STAA
 AC Q07226;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Gamma-hemolysin component B precursor (H-gamma-1) (H-gamma-1).
 DE HLG8 OR SA2209 OR MK2344.
 GN HLG8 OR SA2209 OR MK2344.
 OS Staphylococcus aureus (strain MW2), and
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OK NCBI_TaxID=158879, 196620, 1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N315; 1952; PubMed=11418146;
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaiko C.,
 RA Sekimizu K., Hirakawa H., Kubara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 RT aureus.";
 RT Lancet 357:1225-1240(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MM2;
 RX MEDLINE=22040717; PubMed=12044378;
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
 RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
 RA Yamamoto K., Hiramatsu K.;
 RT "Genome and virulence determinants of high virulence community-
 RT acquired MRSA.";
 RT Lancet 359:1819-1827(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=9313806; PubMed=8423103;
 RX Cooney J.C., Kienle Z., Foster T.J., O'Toole P.W.;
 RT "The gamma-hemolysin locus of Staphylococcus aureus comprises three
 RT linked genes, two of which are identical to the genes for the F and S
 RT components of leukocidin.";


```
RL Infect. Immun. 61:768-771(1993).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89310396; PubMed=3075655;
RA Cooney J.C., Mulvey M., Arbutnot J., Foster T.;
RT "Molecular cloning and genetic analysis of the determinant for gamma-
RT lysin, a two-component toxin of Staphylococcus aureus.";
RT J. Gen. Microbiol. 134:2179-2188(1988).
CC -!- FUNCTION: GAMMA-HEMOLYSIN CAUSES HEMOLYSIS IN RED BLOOD CELLS.
CC -!- SUBUNIT: CONSISTS OF TWO PROTEIN COMPONENTS (H-GAMMA-I AND -II).
CC H-GAMMA-I IS ALSO THE F SUBUNIT OF LEUKOCIDIN.
CC -!- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP003137; BAB43511.1; -.
DR EMBL; AP004830; BAB96209.1; -.
DR EMBL; L01055; AAB26639.1; -.
DR HSP; P09616; 7AHL.
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
DR Hemolysis; Toxin; Signal; Complete proteome.
KW SIGNAL 1 25
FT CHAIN 26 325
FT SEQUENCE 325 AA; 36711 MW; 082999EB9339B6A0 CRC64;
SQ
Query Match 27.2%; Score 404; DB 1; Length 325;
Best Local Similarity 29.4%; Pred. No. 1e-18;
Matches 87; Conservative 74; Mismatches 119; Indels 16; Gaps 6;
QY 9 TGTTDI-GSNTTV---KTGDLVTYDKENG-----MHKKVYFSDIDKNNHKKLLVIR 56
DB 21 SGTANAEKGTTPVSVKVDKVTLYKTATADSKFKISQILTFNFKDKSYDKDTLVLK 80
QY 57 TKTIAQGRVYSEEGANKSLAWPSAFKVLQLPDNEVAQISDYPRNSIDFKYKXSL 116
DB 81 ATGNISGFVKPNPDYDFSKLWGAKNVSISSQNSDVSNNVVDYAPKQNEEFQVNTL 140
QY 117 XYFXGXVGDGXGXIXGXIXAXVIXHXLXYXOPDKTILESPTD-KKVGWKVIFNNMV 175
DB 141 GYTFGGDI--SISNGLSGGLNGTAFSETINYKQESYRTLSRNTYKNVWGVEAHKIM 198
QY 176 QNWGPDYDRSWPNVYGNQLFMKTRNGSKMAADNFDLPNKASSLLSGSPDPFATVITMD 235
DB 199 NNGWGYGRDSFHPYTYGNEFLAGROSSAYAGQNFIAQHOMPLLSRSNFPEFLSVLS-H 257
QY 236 RKASKQNTIDVIYERVRDDYQLHWTSTNWKGTNTKDKWTDSSRSYKIDWEKEEM 291
DB 258 RDGAACKSKITVYQREMDLYQIRWNGFYWAGANYKNFKTRTKSYEIDWENHKV 313
RESULT 3
LUKF_STAAU STANDARD; PRT; 323 AA.
AC P31715;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leukocidin F subunit precursor (Gamma-hemolysin, H-gamma-I subunit).
GN LUKF
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
```

```

RN SEQUENCE FROM N.A., AND SEQUENCE OF 26-66.
RP STRAIN=MRSA NO. 4;
RX MEDLINE=92246945; PubMed=1575738;
RA Rahman A., Nariya H., Izaki K., Kato I., Kamio Y.;
RT "Molecular cloning and nucleotide sequence of leukocidin F-component
RT gene (lukF) from methicillin resistant Staphylococcus aureus.";
RL Biochem. Biophys. Res. Commun. 184:640-646(1992).
RN [2]
RP SEQUENCE OF 27-85 AND 322-323.
RX STRAIN=RIMD 310925;
RX MEDLINE=93223836; PubMed=8467905;
RA Kamio Y., Rahman A., Nariya H., Ozawa T., Izaki K.;
RT "The two Staphylococcal bi-component toxins, leukocidin and gamma-
RT hemolysin, share one component in common.";
RL FEBS Lett. 321:15-18(1993).
CC -!- FUNCTION: LEUKOCIDIN CAUSES CYTOTOXIC CHANGES IN POLYMORPHONUCLEAR
CC LEUKOCYTES. GAMMA-HEMOLYSIN CAUSES HEMOLYSIS IN RED BLOOD CELLS.
CC -!- SUBUNIT: LEUKOCIDIN CONSISTS OF TWO PROTEIN COMPONENTS: F AND S;
CC GAMMA-HEMOLYSIN CONSISTS OF TWO PROTEIN COMPONENTS (H-GAMMA-I-F
CC AND H-GAMMA-II).
CC -!- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S65052; AAC60446.1; -.
DR PIR; JQ1530; JQ1530.
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
DR Hemolysis; Toxin; Signal.
KW SIGNAL 1 25
FT CHAIN 26 323
FT SEQUENCE 323 AA; 36567 MW; A5951CF2AAB6C7E8 CRC64;
SQ
Query Match 26.3%; Score 390; DB 1; Length 323;
Best Local Similarity 28.8%; Pred. No. 7.9e-18;
Matches 85; Conservative 74; Mismatches 120; Indels 16; Gaps 6;
QY 9 TGTTDI-GSNTTV---KTGDLVTYDKENG-----MHKKVYFSDIDKNNHKKLLVIR 56
DB 21 SGTANAEKGTTPVSVKVDKVTLYKTATADSKFKISQILTFNFKDKSYDKDTLVLK 80
QY 57 TKTIAQGRVYSEEGANKSLAWPSAFKVLQLPDNEVAQISDYPRNSIDFKYKXSL 116
DB 81 ATGNISGFVKPNPDYDFSKLWGAKNVSISSQNSDVSNNVVDYAPKQNEEFQVNTL 140
QY 117 XYFXGXVGDGXGXIXGXIXAXVIXHXLXYXOPDKTILESPTD-KKVGWKVIFNNMV 176
DB 141 GYTFGGDI--SISNGLSGGLNGTAFSETINYKQESYRTLSRNTYKNVWGVEAHKIM 198
QY 177 QNWGPDYDRSWPNVYGNQLFMKTRNGSKMAADNFDLPNKASSLLSGSPDPFATVITMDR 236
DB 199 -GWGPGYGRDSFHPYTYGNEFLAGROSSAYAGQNFIAQHOMPLLSRSNFPEFLSVLS-HR 256
QY 237 KASKQNTIDVIYERVRDDYQLHWTSTNWKGTNTKDKWTDSSRSYKIDWEKEEM 291
DB 257 QDRAKSKITVYQREMDLYQIRWNGFYWAGANYKNFKTRTKSYEIDWENHKV 311
RESULT 4
HLGA_STAAU STANDARD; PRT; 309 AA.
ID P31714; Q07225; Q53689; Q53690;
AC
```



```
Db 200 SAYDOYLFAQDPGP-AARDYFVPDNLPLPLOSFGFNPSPITTLISHERGKGRKSE--FEI 256
QY 248 IYERVRDYOQLHWTSTNWKGTNTKDKWTDERS-SERYKIDWEKEEM 291
Db 257 TYGRNMDATYAYVTRHRLAVDRKHDAPFKNRNVTVKYEVNWKTHEV 301

RESULT 5
HLGC_STAUAU STANDARD; PRT; 315 AA.
ID HLGC_STAUAU
AC Q07227;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gamma-hemolysin component C precursor (H-gamma-I) (H-gamma-I).
GN HLGC.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93138806; PubMed=8423103;
RA Cooney J.C., Kienle Z., Foster T.J., O'Toole P.W.;
RT "The gamma-hemolysin locus of Staphylococcus aureus comprises three
RT linked genes, two of which are identical to the genes for the F and S
RT components of leukocidin."
RL Infect. Immun. 61:768-771(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89310396; PubMed=3075635;
RA Cooney J.C., Mulvey M., Arbuthnot J., Foster T.;
RT "Molecular cloning and genetic analysis of the determinant for gamma-
RT lysin, a two-component toxin of Staphylococcus aureus."
RL J. Gen. Microbiol. 134:2179-2188(1988).
CC -1- FUNCTION: GAMMA-HEMOLYSIN CAUSES HEMOLYSIS IN RED BLOOD CELLS.
CC H-GAMMA-I IS ALSO THE F SUBUNIT OF LEUKOCIDIN.
CC -1- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L01055; AAA26638.1; -
CC InterPro; IPR001776; Aerolysin.
CC InterPro; IPR001340; Hemolysn_pore.
CC Pfam; PF01117; Aerolysin; 1.
CC PRINTS; PR01468; BICOMPNTOXIN.
CC TIGRFAMS; TIGR01002; hlyII; 1.
CC Hemolysis; Toxin; Signal.
CC SIGNAL 1 29
CC CHAIN 30 315
CC SEQUENCE 315 AA; 35614 MW; 77359819736620BC CRC64;
CC -----
CC Query Match 16.4%; Score 244; DB 1; Length 315;
CC Best Local Similarity 21.2%; Pred. No. 2e-08;
CC Matches 62; Conservative 84; Mismatches 119; Indels 28; Gaps 10;

QY 11 TTDIGSTNTVKTGDLV--TYDKEN---GMHKKVFYSFIDKHNKLLVIRTKGTAGQY 65
Db 33 TEDIGKGSDD---EIKRTEDKTSNKGVTQNIQDFVCKDKYKNDALILKMQGFISRT 89
QY 66 RVYSEEGANK-SGLAWPSAFKVLQLPDNEVAOISYPRNSIDTKYXSXLXYFXGXV 124
Db 90 TYNYKKTNVHKMRWPFQNYIGLKTNDKYSVLI-NYLPKNKLESTNVSTGLINGGNF 148
QY 125 XGDXGXGXIXAXVIXHXLXYXQPDFKTXILESPSTDKKVGNKVFNNMNVNQNWGPYDR 184
```

```
Db 149 QSAPSLGGNGSF-----NYSKSIYTOQNYVSEVEQONKSVLMGVKANSFATES----- 198
QY 185 DSNPNVYGNOFLMKTRNGSMKADNLDPNKASSLLSSGSPDFATVITMDRKASKOOTN 244
Db 199 -GOKSAFDSLFVGYKPHSKDPRDYFVPDSELPPLVOSGFNPFSFIATVSHE-KGSSTSE 256
QY 245 IDVIYERVRD-----DYQLHWTSTNWKGTNTKDKWTDERS-SERYKIDWEKEEM 291
Db 257 FEITYGRNMDVTHAIKRSTHYGNSYLDGHRVHNAFVNRNRYTVKYEVNWKTHEI 309

RESULT 6
LUKS_STAUAU STANDARD; PRT; 315 AA.
ID LUKS_STAUAU
AC P31716;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leukocidin S subunit precursor.
GN LUKS.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-79.
RC STRAIN=NO.4;
RX MEDLINE=92068180; PubMed=1958181;
RA Rahman A., Izaki K., Kato I., Kamio Y.;
RT "Nucleotide sequence of leukocidin S-component gene (luks) from
RT methicillin resistant Staphylococcus aureus."
RL Biochem. Biophys. Res. Commun. 181:138-144(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93372503; PubMed=7763998;
RA Rahman A., Izaki K., Kamio Y.;
RT "Gamma-hemolysin genes in the same family with lukF and luks genes in
RT methicillin resistant Staphylococcus aureus."
RL Biosci. Biotechnol. Biochem. 57:1234-1236(1993).
CC -1- FUNCTION: LEUKOCIDIN CAUSES CYTOTOXIC CHANGES IN POLYMORPHONUCLEAR
CC LEUKOCYTES.
CC -1- SUBUNIT: LEUKOCIDIN CONSISTS OF TWO PROTEIN COMPONENTS: F AND S.
CC -1- SIMILARITY: TO THE H-GAMMA II SUBUNIT OF GAMMA-HEMOLYSIN.
CC -1- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M81346; AAA26654.1; -
CC DR EMBL; S65052; AAC60445.1; -
CC PIR; JN0626; JN0626
CC InterPro; IPR001776; Aerolysin.
CC InterPro; IPR001340; Hemolysn_pore.
CC Pfam; PF01117; Aerolysin; 1.
CC PRINTS; PR01468; BICOMPNTOXIN.
CC TIGRFAMS; TIGR01002; hlyII; 1.
CC Hemolysis; Toxin; Signal.
CC SIGNAL 1 29
CC CHAIN 30 315
CC SEQUENCE 315 AA; 35557 MW; 6F480166DF13AA1E CRC64;
CC -----
CC Query Match 16.0%; Score 238; DB 1; Length 315;
CC Best Local Similarity 20.8%; Pred. No. 4.8e-08;
CC Matches 61; Conservative 84; Mismatches 120; Indels 28; Gaps 10;

QY 11 TTDIGSTNTVKTGDLV--TYDKEN---GMHKKVFYSFIDKHNKLLVIRTKGTAGQY 65
```

```
Db 33 TEDIGKSDI---EIKRTEDKTSKNGWYQNIQFVQDTRKYNKDALLKMQGFSSRT 89
QY 66 RYSEEGANK-SGLWPSAFKVLQLPDNEVAQISDYPRNSIDTKXXYSXLYXFXGXV 124
Db 90 TYINYKTHVAMRWFOYNGITGLTNDKYVSLI-NYLPKNKIESTNSQTLGYNIGNF 148
QY 125 XGXDGXIXGXIXAXXIXHXIXXQDPFKTITLESPTDKKVGKVFNNQNGPYDR 184
Db 149 QSAPSLGGNGSF----NYSKSISYQOQNVSEVEQNSKSVLWGYKANSFATES----- 198
QY 185 DSWNPVYGQLPFKMTRNGSKMAADNPLDPNKASSLLSSGFSDFATVITMDRKASKQQTN 244
Db 199 -GQSAFSDSLEFGVGYKPHSKDRDYFVDPSELPLVQSGFNPSFIATVSHE-KGSSDTSE 256
QY 245 IDVIVERVDR----DYQLHWTSTNWKGTNTKDKWTDTRS-SERYKIDWEKEEM 291
Db 257 FEITYGRNMDVTHAKRSTHYGNSYLDGHRVHNAFVNRNITVYEVNWKTHIEI 309

RESULT 7
YMBL YEAST
ID YMBL YEAST STANDARD; PRT; 1562 AA.
AC Q04781; Q04029;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 180.2 kDa protein in FAA4-HOR7 intergenic region.
GN YMR247C OR YH9408.09C OR YH9920.01C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE OF 1-956 FROM N.A.
RC STRAIN=S288c / AB972;
RA Hunt S., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 950-1562 FROM N.A.
RC STRAIN=S288c / AB972;
RA Gentles S., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -!- SIMILARITY: STRONG, TO HUMAN ZNF294.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 248639; CAA88574.1; -
DR EMBL; 248756; CAA88657.1; -
DR SGD; S0004861; YMR247C.
DR InterPro; IPR001841; Znf_ring.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; 2F_RING_1; FALSE_NEG.
DR PROSITE; PS50089; 2F_RING_2; 1.
KW Hypothetical protein: zinc-finger.
FT ZN_FING 1508 1555 RING-TYPE.
SQ SEQUENCE 1562 AA; 180185 MW; 97AC65E881362305 CRC64;

Query Match 8.1%; Score 121; DB 1; Length 1562;
Best Local Similarity 19.3%; Pred. No. 21;
Matches 62; Conservative 74; Mismatches 106; Indels 80; Gaps 16;

QY 30 KENGHHKVFYSFIDDK-----NUNKKL-----LVTRTGTIAGQY--RVYSEEGANK 75
Db 1166 KLESQYKRIFEVLNDKDGNSRLLTLLGSLVVKYQQDIIIEYELRIQKQTGSDV 1225
```

```
QY 76 SGLWPSAFKVLQLPDNEVAQISDYPRNSIDTKXXYSXLYXFXGXV----- 124
Db 1226 DGSASDNDVNSFKLPQLKLVQTDVPEKPEYLEYENKNSFIKYLAWHLLMYFKDTSYN 1285
QY 125 -----XGXDGXIXI---XGXIAXVXIXHXIXXQDPFKTITLESPTDKKVGKVFNNM 174
Db 1286 MRQIFIEQLKEAGLINRMFDFTDQIDLRDTEFWQVQDVKETSE-----YNI 1332
QY 175 VNQNGGPDYDROSW---NPVYGQLPFKMTRNGS-----MKAADNPLDPNKASSLLSSGF 224
Db 1333 VGNFSPYKEDIFECKKLLGLTLYQLFNNVNGCLTSIWNLIKDKRTQLQ-NDIEKEFVSFEI 1391
QY 225 SP-----DFATVIT-MDRKASKQQT---NIDVIYERV---DDYQLHWT---STNWK 267
Db 1392 SPILIKNEFDDINSKMDRLTSNDALTIKLNITNEVKASYLIDDOKLEISFKLPKNYPL 1451
QY 268 TNTKDKWTDNR---SSERYKIDW 286
Db 1452 TNIOGVGVSRIQSGQKWK-QW 1472

RESULT 8
CHD3_CAEEL
ID CHD3_CAEEL STANDARD; PRT; 1787 AA.
AC Q22516; Q18794;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chromodomain helicase-DNA-binding protein 3 homolog (CHD-3).
GN CHD-3 OR T1468.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20530482; PubMed=11076750;
RA von Zelewsky T., Palladino F., Brunschwig K., Tobler H., Hajnal A.,
RA Mueller F.;
RT "The C. elegans Mi-2 chromatin-remodelling proteins function in vulval
RT cell fate determination.";
RL Development 127:5277-5284(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol M2;
RA Mathews P., Murray A.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Chromatin-remodelling protein that function in vulval
CC cell fate determination.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
CC -!- SIMILARITY: CONTAINS 2 PHD-TYPE ZINC FINGERS.
CC -!- SIMILARITY: CONTAINS 2 CHROMO DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF308444; AAC29837.1; -
DR EMBL; 267884; CAA91810.1; -
DR EMBL; 267881; CAA91810.1; JOINED.
DR EMBL; 267881; CAA91798.1; -
DR EMBL; 267884; CAA91798.1; JOINED.
DR WormPep; T1468.1; CE03657.
DR InterPro; IPR000953; Chromo.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002464; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR000330; SNF2_N.
```

```

DR InterPro; IPR001965; Znf_PHD.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00176; SNF2_N.1.
DR Pfam; PF00271; helicase_C.1.
DR Pfam; PF00365; chromo.1.
DR Pfam; PF00628; PHD.2.
DR SMART; SM00298; CHROMO.2.
DR SMART; SM00487; DEXDC.1.
DR SMART; SM00490; HELICC.1.
DR SMART; SM00249; PHD.2.
DR SMART; SM00184; RING.2.
DR PROSITE; PS00598; CHROMO.1; FALSE_NEG.
DR PROSITE; PS0013; CHROMO.2.1.
DR PROSITE; PS00690; DEAH_ATP_HELICASE.1.
DR PROSITE; PS01359; ZF_PHD.1.2.
DR PROSITE; PS0016; ZF_PHD.2.2.
KW Chromatin regulator; Nuclear protein; Repeat; Helicase; DNA-binding;
KW ATP-binding; Zinc-finger.
FT DOMAIN 59 62 POLY-LYS.
FT ZN_FING 265 312 PHD-TYPE 1.
FT ZN_FING 328 375 PHD-TYPE 2.
FT DOMAIN 373 476 CHROMO 1.
FT DOMAIN 501 583 CHROMO 2.
FT DOMAIN 1287 1291 POLY-ARG.
FT NP_BIND 641 648 ATP (POTENTIAL).
FT SITE 763 766 DEAH BOX.
SQ SEQUENCE 1787 AA; 205254 MW; 1EFCE1FECE59740 CRC64;

Query Match 7.6%; Score 113.5; DB 1; Length 1787;
Best Local Similarity 17.6%; Pred. No. 79;
Matches 63; Conservative 69; Mismatches 129; Indels 97; Gaps 15;

QY 7 IGTGTTDIGNTVTKGDLVTDYDK-----ENGHMKHVKVYS-----FIDDKHNHK 50
DB 925 LKNGMYE-GSALIKNAGKVFLLQKMLKLDGGHVRVLFQMTMMLDILEDVDFGVGYKY 93
QY 51 KLIVRTKGTIAGYRV-----YSEEGANK-----SGLAWPSAFKVLQLPNEVA 96
DB 984 E-----RIDGSITGQORDAIDRYNAPGAKQVFLSTRAGGLG-----INLATADTVII 1033
QY 97 QISDYPRNSIDP-----KKYKSLXYXF-----XCXVXGDX 129
DB 1034 YSDWNPNDIQAFSRAHRLGQKHVKVMYRFVTKGVSVEERITSVAKKMLLTHLVVRAGL 1093
QY 130 GXIXGXIXAXVHXLYXQPD-FK-----TILESPDCKKVGWK-----V 169
DB 1094 GAKDGKSMSTELDDVLRWGTEELFREEEAPVEGADGEGTSKKPNEQEIWDDAAVDFL 1153
QY 170 IFNNMVNQNGPYDRDSWNPVYGNQLFMKTRNGSMKAADNFDLPNKASSLSGFG---SP 226
DB 1154 LDRNKEEGQDGKKEHWTNXYLSFSFKVATYN--TREADDADDEDETEVIKGTEDQDP 1211
QY 227 DFATVTIMDRKASKQQTNIQIVYVRDYYOLHWTSN-----WKGNTKDKWTDGRSE 280
DB 1212 NYWEKLLKHYYEQDETQELQKLGKVRQVYASENMGQDSANQNOOEEEDDGE 1269

RESULT 9
Y943_METJA
ID Y943_METJA STANDARD; PRT; 417 AA.
AC Q58353;
DT 13-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0943.
GN MJ0943.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;

```

```

RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams R.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.F., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.G., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67538; AAB98953.1; -.
DR TIGR; MJ0943; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 148 168 POTENTIAL.
FT DOMAIN 399 415 PRO-RICH.
SQ SEQUENCE 417 AA; 46355 MW; 6DF8D82E1AA8BD90 CRC64;

Query Match 7.6%; Score 113; DB 1; Length 417;
Best Local Similarity 19.9%; Pred. No. 87;
Matches 56; Conservative 64; Mismatches 113; Indels 48; Gaps 12;

QY 4 DINIKGTDTGNTVTKGDLVTDYDKENGHMKHVKVYSFIDDKHNHKKLLVIRTKGTIAG 63
DB 34 DININTNN-GENTE-KPINQENNVNKNKESQTONIQSYENKEIKNOENHPLOS 91
QY 64 QYRV-----YSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTKYKXSLX 117
DB 92 NQNYEQTNNGFNENENAMTNVGESEVNNPEAYNYIEIT--YPDGTIPDKIEQMLY 149
QY 118 Y-----XFXGXVXGD---XGXIXGXIXAXVHXLYXQPDFFKFTILESPDCKKVGWK 169
DB 150 YIKVIDPIVGLAGDIYVDGNYIGTLDVYGVGECVFY-EPGYHTITAEADNGKILASKT 208
QY 170 I-----FNNMVNQNGPYDRDSWNPVYGNQLFMKTRNGSMKAADNFDLPNKASSLL-- 220
DB 209 VYVEEGTAYNSGESENDEYDNN-----YESNDLQQTQT-QFSEIEVYVDDIRPSNIII 262
QY 221 -----SSGFSPPDFATVITMD--RKASKQQTNIQV 247
DB 263 TKLAMPGLFASINGISPDIGVNIEMENGKINKLYVSMV 303

RESULT 10
AMY_BUTFI
ID AMY_BUTFI STANDARD; PRT; 976 AA.
AC P30269;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase).
GN AMYA.
OS Butyrivibrio fibrisolvens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae;
OC Butyrivibrio.
OX NCBI_TaxID=831;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HI7C;
RX MEDLINE=91286207; PubMed=2061294;

```

RA Rumbak E., Rawlings D.E., Lindsey G.G., Woods D.R.;
 RT "Cloning, nucleotide sequence, and enzymatic characterization of an
 RT alpha-amylase from the ruminal bacterium Butyrivibrio fibrisolvens
 RT H17c.";
 RL J. Bacteriol. 173:4203-4211(1991).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M62507; AAA23005.1; -;
 DR PIR; A42466; A42466.
 DR HSP; P00691; IBAG.
 DR InterPro; IPR000461; Alpha_amylase.
 DR InterPro; IPR005085; CBM_25.
 DR Pfam; PF00128; alpha-amylase; 1.
 DR Pfam; PF02806; alpha-amylase; 1.
 DR Pfam; PF03423; CBM_25; 1.
 KW Hydrolyase; Glycosidase; Carbohydrate metabolism; Signal.
 FT SIGNAL 1 33 POTENTIAL.
 FT CHAIN 34 976 ALPHA-AMYLASE.
 FT ACT_SITE 323 323 BY SIMILARITY.
 FT ACT_SITE 327 327 BY SIMILARITY.
 FT ACT_SITE 447 447 BY SIMILARITY.
 SQ SEQUENCE 976 AA; 106695 MW; 61D690FDC19315AC CRC64;
 Query Match 7.5%; Score 112; DB 1; Length 976;
 Best Local Similarity 20.7%; Pred. No. 38;
 Matches 56; Conservative 47; Mismatches 75; Indels 92; Gaps 15;
 QY 10 GTTDTGNTVTKTGLVYDENGKMKVYFDDKNNKLLVIRKGTAGQYRVYS 69
 DB 276 GTTDLRLTYTSMGLPDPDVTENTGQYFVEFLK-----CVYL 315
 QY 70 EGGANKSGLAWPSAFK---QLQLPDNEVAQISDY-----YPR-----NSID 108
 DB 316 --GAD-----GPRDTAKHISLPDPVP--SDYDAGRNTFFPNMREALNEYSEVG 363
 QY 109 TKXXSXSLXY-----FXGXVXGXDXG-----XIXGXIXAXVXIXHXLX 147
 DB 364 TKSYDELTVYGVLGQTNDRLLAAVQYIGGTTASNYGSSLSLSSGSLNRLLDYQI- 422
 QY 148 YQDPDFKILSPDKVGVKVIENNMYNQN---WGPYDRD-----YGNOLF 196
 DB 423 YDDTAYGSTYATDEKLTWYVESHNDYMNDESECKSIDDDWIMGWSIIAARDAGTPLF 482
 QY 197 MKTRNGSKMAADFELDPNKAASSLLSSGFSFSP 226
 DB 483 FSRPNNS--SAENPYGDN----LIGAGSP 506
 RESULT 11
 ID_CYSF_PLAFA STANDARD; PRT; 569 AA.
 AC P25805;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trophozoite cysteine proteinase precursor (EC 3.4.22.-) (TCP).
 OS Plasmodium falciparum
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92228005; PubMed=1565129;

RA Rosenthal P.J., Nelson R.G.;
 RT "Isolation and characterization of a cysteine proteinase gene of
 RT Plasmodium falciparum.";
 RL Mol. Biochem. Parasitol. 51:143-152(1992).
 CC -1- FUNCTION: PROBABLY DEGRADATES ERYTHROCYTE HEMOGLOBIN.
 CC -1- DEVELOPMENTAL STAGE: TROPHOZOITE
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M81341; AAA29578.1; -;
 DR PIR; A45624; A45624.
 DR HSP; P07686; IQDQ.
 DR MEROPS; C01.077; -;
 DR InterPro; IPR000668; Peptidase_C1.
 DR InterPro; IPR000169; SHprot_acsite.
 DR Pfam; PF00112; Peptidase_C1; 1.
 DR PRINTS; PR00705; PAPAIN.
 DR ProDom; PD000158; Peptidase_C1; 1.
 DR PROSITE; PS00139; THIOLEPROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOLEPROTEASE_HIS; 1.
 DR PROSITE; PS00640; THIOLEPROTEASE_ASN; 1.
 KW Hydrolase; Thiol protease; Zymogen; Glycoprotein; Signal.
 FT SIGNAL 1 ? POTENTIAL.
 FT PROPEP ? 332 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 333 569 TROPHOZOITE CYSTEINE PROTEINASE.
 FT DOMAIN 64 70 POLY-ASN.
 FT ACT_SITE 357 357 BY SIMILARITY.
 FT ACT_SITE 488 488 BY SIMILARITY.
 FT ACT_SITE 533 533 BY SIMILARITY.
 FT DISULFID 354 395 BY SIMILARITY.
 FT CARBOHYD 58 58 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 479 479 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 487 487 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 569 AA; 66880 MW; E5E762A3637F7C60 CRC64;
 Query Match 7.5%; Score 111.5; DB 1; Length 569;
 Best Local Similarity 17.7%; Pred. No. 18;
 Matches 69; Conservative 64; Mismatches 99; Indels 157; Gaps 17;
 QY 10 GTTDTGNTVTKTGLVYDENGKMKVYFDDKNNKLLVIRKGTAGQYRVYS 69
 DB 198 GALEIGMNEEMK-----YRKEDPINNIKYASKFFKFMKEHNKYKNIIDQMRFKFI 251
 QY 41 SFIDDKHNK--KLLVIRKGTAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQ- 97
 DB 252 NYSIKHNKLNKNMYKKK---VNQFSYSEELKE-----YFKTLHPVNHMEKY 301
 QY 98 -----ISDYPRNSIDTKXXYS---XLYFXGXVXGXDXGXIXGXIXAXV 140
 DB 302 SKPFENHLKNILISEFTYNGKRNKEDIKVPILDYREKGIHVHEPKDQGLGCSWAF 361
 QY 141 XIXHLYXQDPFKTILSPDKVGVKVIENNMYNQN---WGPYDRD-----YGNOLF 196
 DB 362 SVGNIESVFAKKNKNIILSFSEQEVV-----DCSKDNFG---CDGHPFYSFLVQLNE 411
 QY 192 ---GNOLFMTNRN-----GSMKAADNFDLPNKAASSL--LSSGFSDFDA 229
 DB 412 LCLGDEYKAKDDMFCLNRYCRKRVKSLSSIGAVKENQILALNEVGLSVNMGVNDVF 471
 QY 230 T-----VITMDRKASKQQTNDIVYERVR 253
 DB 472 AYSEGVYNGTCSEELNHSVLLVGYGVEKTKLNNKNIQTNTKENSQPDNIIY---- 527

```
QY 254 DQYLHWTSN-WKGTNTKDKWDRSSR 281
Db 528 -----YWIKNWS-----SKKGCNGFMR 546

RESULT 12
HX2_HAEIN STANDARD; PRT; 928 AA.
AC P45354;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE Heme/hemopexin-binding protein precursor (Heme:hemopexin utilization
protein A).
GN HXUA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 22-35.
RC STRAIN=DL42 / Serotype B;
RX MEDLINE=95115556; PubMed=7815944;
RA Cope L.D., Thomas S.E., Latimer J.L., Slaughter C.A.,
Mueller-Eberhard U., Hansen E.J.;
RT "The 100 kDa haem:haemopexin-binding protein of Haemophilus
influenzae: structure and localization.";
RL Mol. Microbiol. 13:863-873(1994).
RN [2]
RP SEQUENCE OF 1-30 FROM N.A.
RC STRAIN=DL42 / Serotype B;
RX MEDLINE=95270579; PubMed=7751272;
RA Cope L.D., Yogev R., Mueller-Eberhard U., Hansen E.J.;
RT "A gene cluster involved in the utilization of both free heme and
heme:hemopexin by Haemophilus influenzae type b.";
RL J. Bacteriol. 177:2644-2653(1995).
CC -|- FUNCTION: BINDS HEME/HEMOPEXIN COMPLEXES.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U08348; AAA74138.1; -.
KW Transport; Signal; Repeat.
FT SIGNAL 1 21
FT CHAIN 22 928 HEME/HEMOPEXIN-BINDING PROTEIN.
FT DOMAIN 101 679 6 X 6 AA APPROXIMATE REPEATS.
FT REPEAT 101 106 1-1.
FT REPEAT 205 210 1-2.
FT REPEAT 279 284 1-3.
FT REPEAT 410 415 1-4.
FT REPEAT 635 640 1-5.
FT REPEAT 674 679 1-6.
FT DOMAIN 149 172 4 X 6 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 149 154 2-1.
FT REPEAT 155 160 2-2.
FT REPEAT 161 166 2-3.
FT REPEAT 167 172 2-4.
SQ SEQUENCE 928 AA; 101228 MW; 67D45466A4B92390 CRC64;

Query Match 7.5%; Score 111.5; DB 1; Length 928;
Best Local Similarity 19.7%; Pred. No. 38;
Matches 65; Conservative 70; Mismatches 104; Indels 91; Gaps 15;

QY 1 ADSINIKTGTDTIGSNNTVTKTGLVTDKENG-----MHKKVYFSFDDK 46
Db 379 ADIEDNTGTGTGTGTSSTFTGTGALLANNGKVLKGNVNISGRH-----IDSF 431

QY 47 NHNKLIVIRTKTIAGQYRVYSEEGANKSGLAWPSA-----FKVOLQLPDNEVAQISD 100
Db 432 RGSDDLKLTNKGHI-----DINNADIHSKGRLEFFITSLQNEEDFKSNITITDSKI----- 482
QY 101 YYPNN-----SIDTKKXYS-----XLYYFXGXVXGXDXGXIXGXIXA-----X 139
Db 483 ----NLNGANGMGLGRSVDKDYDNRWQKTEGSKRKFDFKMSVFEFNQVDDVILAGGFEK 538
QY 140 VTXHXLYXQDPDKT-----ILSPDTDKVGVKVFNNMNVNONWGPY--- 182
Db 539 VNLDKIVATGQTFNYIDGVSRRGRKRYEYGVLD--LDKRTQLSEL--NQORRWGYYIDL 594
QY 183 DRDSWNP-VYGNOLFMTKTRNGSKAADNFDLPNKASLLSSGFSFPDFATVITMDRKASKQ 241
Db 595 ELDNMRAYLYRFDLFATKNTGRSTIKDTEINISNINLANKGVHLLAEIKLD----- 648
QY 242 QTNIDVIYERVDYQLHWTSTNWKGTNTK 271
Db 649 NSKIDITFD--KNSQDTLAQTNRGLMNGK 676

RESULT 13
EPIP_STAEP STANDARD; PRT; 461 AA.
AC P30199;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epidermin leader peptide processing serine protease EPIP precursor
DE (EC 3.4.21.-).
GN EPIP.
OS Staphylococcus epidermidis.
OG Plasmid pTu 32.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TU 3298 / DSM 3095;
RX MEDLINE=92155237; PubMed=1740156;
RA Schnell N., Engelke G., Augustin J., Rosenstein R., Ungermann V.,
Goetz F., Entian K.-D.;
RT "Analysis of genes involved in the biosynthesis of lantibiotic
epidermin.";
RL Eur. J. Biochem. 204:57-68(1992).
CC -|- FUNCTION: PROTEASE WHICH CLEAVES THE MATURED LANTIBIOTIC FROM THE
CC MODIFIED PREPEPTIDE (PROBABLE).
CC -|- PATHWAY: Epidermin biosynthesis; last step.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X62386; CAA44257.1; -.
DR PIR; S23420; S23420.
DR HSSP; P00782; 2SBT.
DR MEROPS; S08.060; -.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Serine protease; Signal; Zymogen; Plasmid.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 ? POTENTIAL.
FT CHAIN ? 461 EPIDERMIN LEADER PEPTIDE PROCESSING
FT FT SERINE PROTEASE EPIP.
FT ACT_SITE 149 149 CHARGE RELAY SYSTEM (BY SIMILARITY).
```



```
FT ACT_SITE 194 194 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 402 402 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 461 AA; 51814 MW; F2770F4F8436D906 CRC64;

Query Match 7.5%; Score 111; DB 1; Length 461;
Best Local Similarity 20.9%; Pred. NO. 14;
Matches 60; Conservative 61; Mismatches 120; Indels 46; Gaps 12;

QY 4 DINIKTGTIDGS-----NTVKTGDLVYDKENGHHKVFYFDDKKNHKKLLV--IR 56
DB 149 DSGVNSHTDLKINKIVNEPKNGFRGSEDESGN-----NFEEDKLNHGTLVAGQG 203
QY 57 TRGTIAG-----QYRVYSEGANKGLAWPSAFKVLQQLPDNEVAQIS--DYPRNSIDT 109
DB 204 ANGNLKGVNPGVMVYRVFGSKSEMLVWSKGIIDAANDNDVNVSLGNVLIKDNQNK 263
QY 110 KYYXSLXYFXGVXGXDXGXIXXAXVIXHXL-----XYXQPDFKTILESPTDKYVG 166
DB 264 KKLRRDEKVDYDALQKAINVYAGKGSIVVAAVGNDSINVKKVEIKNRNLNSKTSKV- 322
QY 167 WKVIENNMYNQ-----WGPYDRDSNPV---YGNQLF-MKTRNGSKMAADNFDLPNK 218
DB 323 -----YDSPANLNWTVGSIDNDYISEFSNFGNFDLMTIGGSYK-----LLDKYGGKDA 374
QY 219 LLSGFSFPDFATVITMDRKASKQQTNDIVYERVR-----DDYQL 258
DB 375 WLEKGYMQKQSVLSTSSNGRYIQSGTSLAAPKVSALALEIDKYQL 421

RESULT 14
YCF1_ARATH
ID YCF1_ARATH STANDARD; PRT; 1786 AA.
AC P56785;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 213.7 kDa protein ycf1.
OS Arabidopsis thaliana (Mouse-ear cress).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20039611; PubMed=10574454;
RA Sato S., Nakamura Y., Kaneko T., Asamizu E., Tabata S.;
RT "Complete structure of the chloroplast genome of Arabidopsis
thaliana.";
RL DNA Res. 6:283-290(1999).
CC -!- FUNCTION: NOT YET KNOWN.
CC -!- SIMILARITY: BELONGS TO THE YCF1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP000423; BAA84445.1; -.
CC DR EMBL; AP000423; BAA84433.1; -.
KW Chloroplast; Hypothetical protein.
SQ SEQUENCE 1786 AA; 213727 MW; CFFD2A4D776D7E5D CRC64;

Query Match 7.3%; Score 108.5; DB 1; Length 1786;
Best Local Similarity 20.4%; Pred. No. 1.6e+02;
Matches 68; Conservative 61; Mismatches 110; Indels 95; Gaps 16;

QY 5 INIKTGTIDGSNTVTKGDLVY-DKENGHHK---VFYSFIDDKNNHKKLLVIRTKGT 60
```

```
Db 1092 INPRITTLQFPSTKILDKYIYNEENGKKNTLFISTIKNLISNKKM----- 1144
QY 61 IAGQYRVYSEGANKGLAWPSAFKVLQQLPDNEVAQISDYPRNSIDTKY---XSIX 117
Db 1145 ---SYDLCS-----LSQAYVFYKLSQIKVSNFCKLKAIVLE 1176
QY 118 YFXCXVVGXDXGXIXGXIXAXVIXHXLXXYXQPDFKTILESPTDKYGVW----- 167
Db 1177 YNIC-----ITSFFVKNKIKVFQ-EGHIFHYELNKTFLNSEVQWKNLRSOYQYNLP 1230
QY 168 VYIENNMYNQW-GPYDRDS--WNP-----VYGNQLFEMKTRNGSKMAADNFDLPNK 215
Db 1231 QISWRLVTONWKNKINDSLVNSPLTKEDSYEKKKEDYKKQKFFEADSLNPKHNK 1290
QY 216 -----ASSLSSGFSFPDFATVITMDR---KASKQQTNDIVYERVRDDYQLHWTST 263
Db 1291 KDSIYNLFYKSIHSTEFKNFMSIGIALDNCILVSVFLEKYNIRGMGE-IRHRKYLDWRIL 1349
QY 264 N-----WKGNTNKKDKWTDSSERY-KID 285
Db 1350 NFWFTKVTIETFPWDTKSKKIYINTKVQYQKID 1383

RESULT 15
FAB2_MYCLE
ID FAB2_MYCLE STANDARD; PRT; 420 AA.
AC O69473;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 3-oxoacyl-[acyl-carrier-protein] synthase 2 (EC 2.3.1.41) (Beta-
DE ketoacyl-ACP synthase 2) (KAS 2).
GN KASB OR MLI656 OR MLCBI243.19C.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteriales; Mycobacteriaceae; Mycobacterium.
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -!- FUNCTION: CATALYZES THE CONDENSATION REACTION OF FATTY ACID
CC SYNTHESIS BY THE ADDITION TO AN ACYL ACCEPTOR OF TWO CARBONS
CC FROM MALONYL-ACP (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + CO(2) +
CC [acyl-carrier protein].
CC -!- PATHWAY: Fatty acid biosynthesis. Involved in meromycolate
CC extension.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: BELONGS TO THE BETA-KETOACYL-ACP SYNTHASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL023635; CAA19200.1; ALT_INIT.
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:40:04 ; Search time 28.5 seconds
(without alignments)
2118.309 Million cell updates/sec

Title: CLAIM6

Perfect score: 1485

Sequence: 1 ADSINIKTGTDTGSNTTV.....WDRSSRYKIDWEKEETN 293

Scoring table: BLOSUM62Dx

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1481	99.7	319	16 Q99UU6	Q99uu6 staphylococ
2	459.5	30.9	412	2 Q05387	Q05387 bacillus ce
3	409	27.5	325	2 Q57227	Q57227 staphylococ
4	404	27.2	327	2 Q54082	Q54082 staphylococ
5	403.5	27.2	327	2 Q93UU8	Q93uu8 staphylococ
6	402.5	27.1	327	16 Q99T54	Q99t54 staphylococ
7	402.5	27.1	332	16 Q931N4	Q931n4 staphylococ
8	401	27.0	325	16 Q931F3	Q931f3 staphylococ
9	400	26.9	325	2 Q50604	Q50604 staphylococ
10	399	26.9	338	16 Q99SN8	Q99sn8 staphylococ
11	396	26.7	325	9 Q80067	Q80067 staphylococ
12	394.5	26.6	326	2 Q54327	Q54327 staphylococ
13	393	26.5	325	2 Q53747	Q53747 staphylococ
14	383	25.8	322	2 Q53704	Q53704 staphylococ
15	383	25.8	322	9 Q9MBN2	Q9mbn2 staphylococ
16	364	24.5	336	2 Q9EVAL	Q9eval bacillus ce

17	333	22.4	265	2	Q937V0	Q937v0 bacillus ce
18	332	22.4	265	2	Q937V1	Q937v1 bacillus ce
19	325.5	21.9	260	2	Q937V2	Q937v2 bacillus ce
20	309	20.8	336	2	Q46308	Q46308 clostridium
21	309	20.8	336	2	Q46181	Q46181 clostridium
22	305	20.5	309	2	Q9L403	Q9l403 clostridium
23	260	17.5	308	2	Q53731	Q53731 staphylococ
24	257.5	17.3	351	16	Q931I5	Q931i5 staphylococ
25	257	17.3	308	2	Q53703	Q53703 staphylococ
26	257	17.3	308	9	Q9MBN3	Q9mbn3 staphylococ
27	255.5	17.2	331	16	Q99SN7	Q99sn7 staphylococ
28	247	16.6	315	2	Q53701	Q53701 staphylococ
29	245	16.5	286	2	Q9AFA9	Q9afa9 staphylococ
30	244	16.4	286	2	Q9AFB0	Q9afb0 staphylococ
31	244	16.4	315	2	Q53691	Q53691 staphylococ
32	244	16.4	315	16	Q99RL1	Q99rl1 staphylococ
33	241.5	16.3	311	2	Q93UU9	Q93uu9 staphylococ
34	241.5	16.3	311	16	Q99T53	Q99t53 staphylococ
35	237	16.0	312	2	Q50603	Q50603 staphylococ
36	237	16.0	312	9	Q80066	Q80066 staphylococ
37	233	15.7	314	2	Q54081	Q54081 staphylococ
38	231	15.6	312	2	Q53746	Q53746 staphylococ
39	201.5	13.6	310	2	Q54326	Q54326 staphylococ
40	147	9.9	287	2	P97144	P97144 staphylococ
41	124.5	8.4	217	2	Q93CF5	Q93cf5 lactococcus
42	123.5	8.3	454	16	Q8RGZ8	Q8rgz8 fusbacteri
43	122	8.2	364	2	Q68629	Q68629 haemophilus
44	121	8.1	360	2	Q68626	Q68626 haemophilus
45	121	8.1	572	3	P87247	P87247 botrytis ci

ALIGNMENTS

RESULT 1

Q99UU6 ID Q99UU6 PRELIMINARY; PRT; 319 AA.
AC Q99UU6;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Alpha-hemolysin precursor.
GN SAV1163 OR SA1007.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iano J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Murayama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RA "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003361; BAB57325.1; -
DR EMBL; AP003132; BAB42258.1; -
DR HSSP; P09616; TAAHL.
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMENTOXIN.
DR TIGRfams; TIGR01002; hlyII; 1.
DR PROSITE; PS00274; AEROLYSIN; 1.
KW Complete proteome.

[illegible]

QY	86	VOLQLPDNEVAQISDYPRNSIDTKYXSXLXYFXG--XVXGDXG-XIXGXIXAXVXI	142
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	111	TEAEITSGDSQAQFHAAPVNTMTSAKVTSEVGYTLGGSVKVGVNDKGNADASITGSFAW	170
		: : : : : : : : : : : : : : : : : : : : : : : : :	
QY	143	XHLXLYXOPDFKTTILESPDTKKVGKVKVIFNNMNVNQNGPYDRDSWNPVYGNQLFMKTRNG	202
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	171	KESYSYDQVDYKTVLEHTHTOKLKNWKYGFOSFNPPEWGIYNRDSFNTFYGNQLFMKRSY	230
		: : : : : : : : : : : : : : : : : : : : : : : : :	
QY	203	SMKAADNFLDPNKASSLLSGFSDFATVLTMDRKAASKQQTINIDVIYERVDDYQLHWTS	262
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	231	N-EGTNFNSKDTVPALTGYGFSNVVAVITADK--TESTSDLKITNRRISDQINIEWWS	287
		: : : : : : : : : : : : : : : : : : : : : : : : :	
QY	263	TNWGTTNTKDKWTDRRSERYKIDWEKEEMT	292
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	288	SKWGTNNKDTYNEFFNNYKLDWKNHQVT	317
		: : : : : : : : : : : : : : : : : : : : : : : : :	
RESULT	3		
Q57227	ID	Q57227 PRELIMINARY; PRT; 325 AA.	
AC	Q57227:		
DT	01-NOV-1996 (TReMBLrel. 01, Created)		
DT	01-NOV-1996 (TReMBLrel. 01, Last sequence update)		
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)		
DE	HLGB-like ORF precursor.		
DE	HLGB-LIKE ORF OR LUK F-R.		
GN	Staphylococcus aureus.		
OS	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;		
OC	Staphylococcus.		
OX	NCBI_TaxID=1280;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 49775;		
RX	MEDLINE=96009775; PubMed=7558328;		
RA	Prevost G., Couplet P., Petiau P., Supersac G., Finck-Barbancon V.,		
RA	Montell H., Piemont Y., Gibier B.;		
RT	"Panton-Valentine leucocidin and gamma-hemolysin from Staphylococcus		
RT	aureus ATCC 49775 are encoded by distinct genetic loci and have		
RT	different biological activities.";		
RL	Infect. Immun. 63:4121-4129(1995).		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=P83;		
RA	Prevost G.;		
RL	Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.		
RP	[3]		
RP	SEQUENCE OF 1-186 FROM N.A.		
RC	STRAIN=P83;		
RX	MEDLINE=93138781; PubMed=8423088;		
RA	Supersac G., Prevost G., Piemont Y.;		
RT	"Sequencing of leucocidin R from Staphylococcus aureus P83 suggests		
RT	that staphylococcal leucocidins and gamma-hemolysin are members of a		
RT	single, two-component family of toxins.";		
RL	Infect. Immun. 61:580-587(1993).		
RP	[4]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=P83;		
RA	Prevost G.;		
RL	Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.		
RP	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=P83;		
RA	Prevost G.;		
RL	Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.		
RP	[6]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=P83;		
RX	MEDLINE=95210788; PubMed=7696748;		
RA	Prevost G., Supersac G., Piemont Y., Colin D., Sire S., Henslen T.,		
RA	Petiau P., Meunier O., Gibier B., Koenig W., Piemont Y.;		
RT	"A particular class of virulence factors: calmodulin-activated		
RT	bacterial adenylate cyclases.";		
RL	Zentralbl. Bakteriol. 24:284-293(1994).		

DR	EMBL; X81586; CAA57278.1; -.
DR	EMBL; X64389; CAA45729.1; -.
DR	HSSP; P09616; 7AHL.
DR	InterPro; IPR001776; Aerolysin.
DR	InterPro; IPR001340; Hemlysn_pore.
DR	InterPro; IPR003963; Staph_bicn_txn.
DR	Pfam; PF01117; Aerolysin; 1.
DR	PRINTS; PR01468; BICOMPNTOXIN.
DR	TIGRFAMS; TIGR01002; hlyII; 1.
KW	SIGNAL.
FT	SIGNAL.
FT	CHAIN
FT	26 325
SEQ	SEQUENCE 325 AA; 36812 MW; 38A0066D8DBD8935 CRC64;
 Query Match	
Best Local Similarity 27.5%; Score 409; DB 2; Length 325;	
Matches 81; Conservative 73; Mismatches 112; Indels 16; Gaps	
QY	11 TTDTGSNTVTKTGDLYVDKENGHHKVFFSIDDKNNKKLLVIRTKGTIAGQVRVYSE 70 : : : : : : : : :
Db	47 TTATADSDKFISQILT-----FNFIKDSYDKDTLVLKAAGNINSGYERPNP 94 : : : : : : : : :
QY	71 EGANKSGLAWPFAFKVOLQLPDNEVAOISDYPPNSIDTKXXSXSLXYFXCVXGXDXG 130 : : : : : : : : :
Db	95 KDYDFSKIYGAKYNVSLSOSSDNVVDPAPKNQNEEFQQNLTGYTFGGDI--SISN 152 : : : : : : : : :
QY	131 XIXGXIXAXYHXHXLXYQPDKFTILESPTD-KKVGKVFNNMNQNGPYDRDSWNP 189 : : : : : : : : :
Db	153 GLSGGLGNATAFSETINYKOESYRTILSRNTNYKNVGVGEAHKIMNNGWPGYGRDSFP 212 : : : : : : : : :
QY	190 VYGNLFPMKTRNGSMKAADNFLDPNKASLLSSGFSPDAFVIYMDRKASKOOTNDIVY 249 : : : : : : : : :
Db	213 TYGNELFLAGRQSAYAGQNFIAQHMPLLSRSNFNPEFLSVLS-HRODGAKKSKITVTY 271 : : : : : : : : :
QY	250 ERVDDYQLHWTSNWKTNTKDWTDRSSERYKIDWEKEE 291 : : : : : : : : :
Db	272 QREMDLQIRWNGFYWAGANYKFKTRFKTYELDWNHKV 313 : : : : : : : : :
 RESULT 4	
O54082	PRELIMINARY; PRT; 327 AA.
ID O54082	AC O54082 PRELIMINARY; PRT; 327 AA.
DC AC O54082	01-JUN-1998 (TrEMBLrel. 06, Created)
DT DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)	
DT DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)	
DE DE Leukotoxin, Lukd precursor.	
GN GN LUKD.	
OS OS Staphylococcus aureus.	
OC OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;	
OC OC Staphylococcus.	
ON ON NCBI_TaxID=1280;	
OR OR [1]	
RP RP SEQUENCE FROM N.A.	
RZ RZ STRAIN=NEWMAN;	
RX RX MEDLINE=98452939; Pubmed=9781679;	
RA RA Gravet A., Colin D.A., Keller D., Giradot R., Monteil H., Prevost G.;	
RT RT "Characterization of a novel structural member, Luke-Lukd, of the bi-	
RL RL component staphylococcal leukotoxins family.";	
RL FEBS Lett. 436:202-208(1998).	
DR EMBL; Y13225; CAA73668.1; -.	
DR InterPro; IPR001776; Aerolysin.	
DR InterPro; IPR001340; Hemlysn_pore.	
DR InterPro; IPR003963; Staph_bicn_txn.	
DR Pfam; PF01117; Aerolysin; 1.	
DR PRINTS; PR01468; BICOMPNTOXIN.	
DR TIGRFAMS; TIGR01002; hlyII; 1.	
KW SIGNAL.	
KW FT	
FT CHAIN	
FT 27 327	
SEQ SEQUENCE 327 AA; 36876 MW; 6CC5677521C3CD24 CRC64;	
 Query Match	
Best Local Similarity 27.2%; Score 404; DB 2; Length 327;	


```
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;  
RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
RL aureus".  
DR Lancet 357:1225-1240(2001).  
DR EMBL; AP003365; BAB58583.1; -.  
DR InterPro; IPR001776; Aerolysin.  
DR InterPro; IPR001340; Hemlysn_pore.  
DR Pfam; PF01117; Aerolysin; 1.  
DR TIGRFAMS; TIGR01002; hlyII; 1.  
DR Complete proteome.  
KW SEQUENCE 325 AA; 36658 MW; 082999EB877CA2E5 CRC64;  
SQ  
  
Query Match 27.0%; Score 401; DB 16; Length 325;  
Best Local Similarity 29.4%; Pred. No. 3.7e-18;  
Matches 87; Conservative 74; Mismatches 119; Indels 16; Gaps  
6;  
  
QY 9 TGITDI--GSNTTV---KYGDLVIYDKENG-----MHKKVYFSFIDDKNNHKKLVIR 56  
DB 21 SGTANAEGKITPVSVKKVDDKVTLYKTATADSDKFKISQILTFFNFIDKRSYDKDITLVLK 80  
  
QY 57 TKGTIAGQYRYSSEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTKKXXSL 116  
DB 81 ATGNINSGFVKPNPDYDFSKLYGAKYVNSISSQSDNSVNVVDYAPKNQNEEFQVQNTL 140  
  
QY 117 XYFXGXVXGXDXGXIXGXIXAXVIXHXLXYQDPFKTILESPTD-KKYGWKVIFNNMV 175  
DB 141 GYTEGGDI--SISNGLSGGLNGTAFSETINYKQESYRTSLSRNTYKNVGNWGVGAHKIM 198  
  
QY 176 NQWKGPDYDRSWPNVYGNQLFMKTRNGSMKAADNFPDKNASSLLSGSFGSPDFATVITMD 235  
DB 199 NNGWGPYGRDSFHTPTGYNELFLAGRSSAYAGONFIAQHOMPLLSRNFNFPEFLSVLS-H 257  
  
QY 236 RKASKQOTNIDVIYERYRVDYQLHWTSTNWKGTNTKDKWTRSSERYKIDWEKEEM 291  
DB 258 RQDGAKSKITVTYQREMDLYQICWNGFYWAGANYKNFKRTFKSTYEIDWENHKV 313  
  
RESULT 9  
O50604 PRELIMINARY; PRT; 325 AA.  
ID O50604  
AC O50604  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DE 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE LuKF-pv.  
GN LuKF-pv.  
OS Staphylococcus aureus.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Staphylococcus.  
OX NCBI_TaxID=1280;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC49775;  
RX MEDLINE=98067870; PubMed=9404084;  
RA Kaneko J., Muramoto K., Kamio Y.;  
RT "Gene of LuKF-pv-like component of panton-Valentine leukocidin in  
RL Staphylococcus aureus p83 is linked with lukM.";  
FL Biosci. Biotechnol. Biochem. 61:1960-1962(1997).  
DR EMBL; AB006796; BAA2408.1; -.  
DR InterPro; IPR001776; Aerolysin.  
DR InterPro; IPR001340; Hemlysn_pore.  
DR InterPro; IPR003963; Staph_bicn_txn.  
DR Pfam; PF01117; Aerolysin; 1.  
DR PRINTS; PR01468; BICOMPOTOXIN.  
DR TIGRFAMS; TIGR01002; hlyII; 1.  
SQ SEQUENCE 325 AA; 36958 MW; 19522EA2DA49824 CRC64;  
  
Query Match 26.9%; Score 400; DB 2; Length 325;  
Best Local Similarity 27.3%; Pred. No. 4.3e-18;  
Matches 76; Conservative 77; Mismatches 109; Indels 16; Gaps  
4;  
  
QY 11 TTDIGSNTTVKTDGLVITYDKENGHHKKVYFSFIDDKNNHKKLVIRTKGTIAGQYRVSE 70  
DB 258 RQDGAKSKITVTYQREMDLYQICWNGFYWAGANYKNFKRTFKSTYEIDWENHKV 313  
  
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;  
RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
RL aureus".  
DR Lancet 357:1225-1240(2001).  
DR EMBL; AP003365; BAB58583.1; -.  
DR InterPro; IPR001776; Aerolysin.  
DR InterPro; IPR001340; Hemlysn_pore.  
DR Pfam; PF01117; Aerolysin; 1.  
DR TIGRFAMS; TIGR01002; hlyII; 1.  
DR Complete proteome.  
KW SEQUENCE 325 AA; 36658 MW; 082999EB877CA2E5 CRC64;  
SQ  
  
Query Match 26.9%; Score 399; DB 16; Length 338;  
Best Local Similarity 28.1%; Pred. No. 5.4e-18;  
Matches 81; Conservative 85; Mismatches 110; Indels 12; Gaps  
7;  
  
QY 13 DIGSNTTVKTDGLVITYDKENGHHKKVYFSFIDDKNNHKKLVIRTKGTIAGQYRVSEEG 72  
DB 42 NLDGTRKMYTRTATTSDSQNKNTQSLQFNELTPEYDKETVFIKAKGTISGLRILDPNG 101  
  
QY 73 ANKSGLAWPSAFKVLQ-LPDNEVAQISDYYPNRSIDTKKXXSLXXYFXGXVXGXDXG 131  
DB 102 YWNTSLRWPGSYSVISIQVNDNNNTNTDFAPKNQDSREVKYTYGKTGCD-FSINRGG 160  
  
QY 132 IXGXIXAXVIXHXLXYQDPFKTILESPTD-KKYGWKVIFNNMVNQNNGPYGRDSW 187  
DB 258 RQDGAKSKITVTYQREMDLYQICWNGFYWAGANYKNFKRTFKSTYEIDWENHKV 313
```

```
Db 161 LAGNITKESNYSETISYQQPSYRTLLDQSTSHKGVGWKVEAHLLNNMGHDHTROLTNDSD 220
QY 188 NPVYGNQLEFMKTRNGSMKADNFDLPNKASSLLSSGSPDFATVITMDRKASKQQTNDIV 247
Db 221 NRT-KSEISLFRNGNLWAKDNFTPKDKPVPVYSEGFPFLAVYSHDKK-DKGKSOQVV 278
QY 248 IYERVDYQLHWTSN---WKGNTTKDKWDRSSERYKIDWEKEM 291
Db 279 HYKRSMDEPKIDNRHGFYSGENHVDKKEKLSALYEVWDKTHDV 326

RESULT 11
OB0067 PRELIMINARY; PRT; 325 AA.
AC OB0067;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE LUKF-PV.
GN LUKF-PV.
OS Staphylococcus aureus bacteriophage PVL, and
OS Staphylococcus aureus temperate phage phisLT.
OC Viruses; dsDNA viruses; no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=71366; 130478;
RN [1]
RP SPECIES-Phase phi PVL;
RC MEDLINE=98067870; PubMed=9404084;
RA Kaneko J., Kimura T., Kawakami Y., Tomita T., Kamio Y.;
RT "Panton-valentine leukocidin genes in a phage-like particle isolated
RT from mitomycin C-treated Staphylococcus aureus V8 (ATCC 49775).";
RL Biosci. Biotechnol. Biochem. 61:1960-1962(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-Staphylococcus aureus temperate phage phisLT;
RA Narita S., Kaneko J., Chiba J., Etienne J., Piemont Y., Kamio Y.;
RT "Phage conversion of Panton-Valentine leukocidin (PVL) in
RT Staphylococcus aureus: molecular analysis of a PVL-converting phage,
RT phisLT.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB009866; BAA31901.1; -
DR EMBL; AB045978; BAB21755.1; -
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemolysn.pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
DR SEQUENCE 325 AA; 36962 MW; 1955C10D28DFCC4F CRC64;

Query Match 26.7%; Score 396; DB 9; Length 325;
Best Local Similarity 27.0%; Pred. No. 7.9e-18;
Matches 75; Conservative 77; Mismatches 110; Indels 16; Gaps 4;

QY 11 TTDIGSNTVTKGDLVTDKENGMMKVFYFIDDKNNKLLVIRTKGTAGQYRVYSE 70
Db 45 FTATSDSKLKSQILT-----FNFKDKSYDKDITLILKAAGNIYSYGTKNP 92
QY 71 EGANKSLWPSAFKVLQLPDNEVAQISDYYPNRSIDTKYXSLXVYXGXVXGXDXG 130
Db 93 KDTISSQFYWGSKYINISNSDSNVVDYAPKNQNEEFQVQQTGVSYGDI--NISN 150
QY 131 XIXGXIXAXVIXIXHLXYXQDFKLTILESPTD-KVGVKVFIFNNMNVNMGFYDRDSWNP 189
Db 151 GLSGGGGSKSPSETINYKQESYRTTSLDKRTNFKIGDVEAHKIMNNGWGPYGRDYSHS 210
QY 190 VYGNOLFVKTRNGSMKADNFDLPNKASSLLSSGSPDFATVITMDRKASKQQTNDIVY 249
Db 211 TYGNEWFLGSRQSNLNAQNFLEHKKMPVLSRGNFNPFIQVLSRKQNAAR-KSKITVY 269
QY 250 ERVRDYLQHWTSNWKNTTKDKWDRSSERYKIDWE 287
Db 270 QREMDRYTFNFWNLHWIGNNYKDNENRATHTSIYVDWE 307
```

```
RESULT 12
Q54327 PRELIMINARY; PRT; 326 AA.
AC Q54327;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Synergohymenotropic toxin.
GN LUKF.I.
OS Staphylococcus intermedius.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1285;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B62;
RX MEDLINE=96105366; PubMed=7498527;
RA Prevost G., Bouakham T., Piemont Y., Montell H.;
RT "Characterisation of a synergohymenotropic toxin produced by
RT Staphylococcus intermedius.";
RL FEBS Lett. 376:135-140(1995).
DR EMBL; X79188; CAA55783.1; -
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemolysn.pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
FT CHAIN 27 326 SYNERGOHYMENOTROPIC TOXIN.
SQ SEQUENCE 326 AA; 36597 MW; F8A15CE8345A22B4 CRC64;

Query Match 26.6%; Score 394.5; DB 2; Length 326;
Best Local Similarity 29.6%; Pred. No. 1e-17;
Matches 81; Conservative 76; Mismatches 108; Indels 9; Gaps 8;

QY 18 TTVTKGDLVTVTKEN-GMHKKVFYSFIDDKNNKLLVIRTKGTAGQYRVYSEGANKS 76
Db 43 TLYKT--TATADSDNLNISQLLTFFNFKDKSYDKDITLVKKAAGNINSYKSPNPNDIYS 100
QY 77 GLWPSAFKVLQLPDNEVAQISDYYPNRSIDTKYXSLXVYXGXVXGXDXGXIXXI 136
Db 101 SFYWGAKYNSISAESKGVNVVDYAPKNQNEEFQVQNTLGYSEGGDIS-ISKG-LSGGL 158
QY 137 XAXVIXHLXYXQDFKLTILESPTDKK-VGVKVFIFNNMNVNMGFYDRDSWNPVYGNOL 195
Db 159 NGSESFSETINYKQESYRTTIDKHTDNKTIGWVEAHKIMNAGWGPYGRDSFHDLYGNEL 218
QY 196 FMKTRNGSMKADNFDLPNKASSLLSSG-FSPDFATVITMDRKASKQQTNDIVYERV 254
Db 219 FLGGRQSKLNAGQNF-LPTRQMPLLARGNFNPEFLSVLSHKPNGAK-TSKIKVTVQREMD 276
QY 255 DYQLHWTSNWKNTTKDKWDRSSERYKIDWEK 288
Db 277 EYTNWNGFHWGTYNKNQNNATFTSFYEIDWDQ 310

RESULT 13
Q53747 PRELIMINARY; PRT; 325 AA.
AC Q53747;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Leucocidin F precursor.
GN LUKF.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE OF 1-285 FROM N.A.
```



```
RC STRAIN=ATCC 49775;
RA Prevost G., Supersac G., Piemont Y., Colin D., Sire S., Henslen T.,
RA Petiau P., Meunier O., Gibier B., Koenig W., Piermont Y.,
RT "The new family of leukotoxins from S.aurea: structural and
RT biological properties."; 284-293(1994).
RL Zentralbl. Bakteriol. 24:284-293(1994).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 49775;
RA Prevost G.;
RN Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
[3]
RC SEQUENCE FROM N.A.
RX MEDLINE=96009775; PubMed=7558328;
RA Prevost G., Couppie P., Petiau P., Supersac G., Finck-Barbancon V.,
RA Montell H., Piemont Y., Gibier B.;
RT "Panton-Valentine leukocidin and gamma-hemolysin from
RT staphylococcus aureu ATCC 49775 are encoded by distinct loci and have
RT different biological activities.";
RL Infect. Immun. 63:4121-4129(1995).
DR EMBL: X72700; CAA51252.1; -
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
KW SIGNAL.
FT SIGNAL. 1 24 POTENTIAL.
FT CHAIN 25 325 POTENTIAL.
SQ SEQUENCE 325 AA; 37009 MW; 03C2C247B3B3CCD9 CRC64;

Query Match 26.58; Score 393; DB 2; Length 325;
Best Local Similarity 27.08; Pred. No. 1.2e-17;
Matches 75; Conservative 76; Mismatches 111; Indels 16; Gaps 4;

Qy 11 TTDTGSNTVTKTGDVLTVDKENGHMKKVFYFIDDKHNHKKLVIRTKTAGYRVYSE 70
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 45 TTATSDSKLKSILT-----FNFIDKSYDKDRTLKKAAGNIYSYGYTRPNP 92

Qy 71 EGANKSGLAPSAFKVOLQDPNEVAQISDYPRNSIDTKYXSLXLYFXGXVXGDXG 130
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 93 KDTISSQFYMGSKYKYNISNDSNDSNVVDYAPKNQNEEFQVQTVGYSGGDI--NISN 150

Qy 131 XIXGXIXAXVIXHXLYXQPDFKILESPTD-KKVGKVFIFNNVQNWGPYDRDSWNP 189
::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 GLSGGGNGSKSFSETINYKQESYRTSLDKRTNFKKIGWDVEAHKIMNMGWGPYGRDSYHS 210

Qy 190 VYGNQLFMKTRNGSKAADNFDLPNKASSLSGSPDFATVITMDRKASKQQTNDIVY 249
||: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 211 TYGNEMFLGSQSNLAGQNFLEHKKMPVLSRGNFNFEFGVLSRKQNAK-KSKITVY 269

Qy 250 ERVRDDYQLHWTSTNWKGTNTKDKWTRSSRYKIDWE 287
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 270 QSEMDRYTNFENFENWIGNNYKDHIRATHTSIYEDWE 307

RESULT 14
Q53704 ID Q53704 PRELIMINARY; PRT; 322 AA.
AC Q53704;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE LukF-PV like component.
GN LukF-PV.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=P83;
RA Kaneko J., Choerit W., Kamio Y.;
RT "LukF-PV like component of leukocidin and gamma-hemolysin in S.aureus
RT P8 3.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D83951; BAA12148.1; -
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
SQ SEQUENCE 322 AA; 36519 MW; 9424E58715B108B6 CRC64;

Query Match 25.88; Score 383; DB 2; Length 322;
Best Local Similarity 26.38; Pred. No. 5.6e-17;
Matches 80; Conservative 78; Mismatches 122; Indels 24; Gaps 4;

Qy 6 NIKGTGTDIGSNTVTKTGDVLTVDKENGHMKV-----FVSFIDD 45
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 10 SVATSIILMLSNVDAQAHTTPVSEKKVDDKITLYKTATSDSKLKISQILTFNFKD 69

Qy 46 KHNHKKLVIRTKTAGYRVYSEEGANKSGLAPSAFKVOLQDPNEVAQISDYTPRN 105
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 KSYDKDLILKKAAGNIYSGYTQPTSDSINSQFYWGAKYVNFVSESCKDSYNIYDAPKN 129

Qy 106 SIDTKYXSLXLYFXGXVXGDXGXIXGXIXAXVIXHXLYXQPDFKILESPTD-KK 164
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 QNEEFQVQTVGYSGGDI--NIINGLTGGLNGSKSFSETINYKQESYRTTIDRKTNHS 187

Qy 165 VGVKVFIFNNVQNWGPYDRDSWNPVYCNQLFMKTRNGSKMAADNFDLPNKASSLSGFG 224
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 IGWGEAHKIMNMGWGPYGRDSSDSLYGNELFLGGRSSSNANQNFPTHOMPILARGNF 247

Qy 225 SPDFATVITMDRKASKQQTNDIVYERVDDYQLHWTSTNWKGTNTKDKWTRDSERYKI 284
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 NPEISVLSLHKQDKV-KSKIKVYQREMDRYENFNWNLHIGYNIKNIKRATHTSIVEI 306

Qy 285 DWK 288
||||
Db 307 DWK 310

RESULT 15
Q9MBN2 ID Q9MBN2 PRELIMINARY; PRT; 322 AA.
AC Q9MBN2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE LukF-PV(P83).
OS Staphylococcus aureus prophage phiPV83.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=129009;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P83;
RA Zou D., Kaneko J., Narita S., Kamio Y.;
RT "Complete nucleotide sequence and molecular characterization of
RT prophage PV83pro carrying lukM-lukF-pv(p83) gene cluster in
RT Staphylococcus aureus strain P83.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=P83;
RX MEDLINE=98067870; PubMed=9404084;
RA Kaneko J., Muramoto K., Kamio Y.;
RT "Gene of LukF-PV-like component of Panton-Valentine leukocidin in
RT Staphylococcus aureus P83 is linked with lukM.";
RL Biosci. Biotechnol. Biochem. 61:1960-1962(1997).
DR EMBL; AB044554; BAA97867.1; -
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:38:18 ; Search time 33 Seconds
(without alignments)
1183.104 Million cell updates/sec

Title: CLAIM7

Perfect score: 1484

Sequence: 1 ADSIDINIKTGTIGSNTTV.....WDRSSERYKIDWEKEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_101002:*
- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
 - 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
 - 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
 - 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
 - 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
 - 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
 - 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
 - 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
 - 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
 - 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
 - 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
 - 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
 - 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
 - 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
 - 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
 - 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
 - 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
 - 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
 - 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
 - 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
 - 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
 - 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
 - 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1484	100.0	293	20	AAV01365
2	421.5	28.4	330	22	AAU33797
3	421.5	28.4	333	22	AAU36528
4	395.5	26.7	327	22	AAU34405
5	395.5	26.7	327	22	AAU37136
6	395.5	26.7	327	22	AAU37564
7	388	26.1	325	22	AAU34287
8	388	26.1	325	22	AAU37100
9	388	26.1	325	22	AAU37186
10	388	26.1	325	22	AAU37477

11	388	26.1	325	23	AAU75478	S. aureus antigeni
12	383	25.8	323	22	AAU34395	Staphylococcus aur
13	383	25.8	328	22	AAU37509	Staphylococcus aur
14	374	25.2	323	14	AAU35084	(Pro)leukocidin F.
15	253	17.0	321	22	AAU34275	Staphylococcus aur
16	253	17.0	321	22	AAU37101	Staphylococcus aur
17	253	17.0	321	22	AAU37128	Staphylococcus aur
18	242	16.3	315	14	AAU35083	(Pro)leukocidin S.
19	240	16.2	312	22	AAU33771	Staphylococcus aur
20	133.5	9.0	826	22	AG81986	S. epidermidis ope
21	133.5	9.0	10182	23	ABP38314	Staphylococcus epi
22	132.5	8.9	1300	22	AAU33407	Enterococcus faeca
23	132.5	8.9	1300	22	AAU35313	Enterococcus faeca
24	129	8.7	40	20	AAV01366	Alpha haemolysin (
25	124.5	8.4	1231	20	AAV00219	Enterococcus faeca
26	124.5	8.4	1231	23	ABP43438	E faecalis EF108 a
27	124.5	8.4	1265	20	AAV00218	Enterococcus faeca
28	118	8.0	447	23	ABP43437	E faecalis EF108 p
29	117.5	7.9	2478	22	ABO9643	Mutp protein of th
30	117.5	7.9	2478	22	AAU34320	Staphylococcus aur
31	117.5	7.9	2478	22	AAU37374	Staphylococcus aur
32	117	7.9	978	22	AAU33960	Staphylococcus aur
33	117	7.9	1001	22	AAU37093	Staphylococcus aur
34	116.5	7.9	1249	22	ABG21666	Novel human diagno
35	116	7.8	384	15	AAU47236	Wild-type Feline H
36	116	7.8	435	23	ABP55618	Lactococcus lactis
37	116	7.8	2893	19	AAW98828	H. pylori GHPO 148
38	115.5	7.8	2893	19	AAW71556	Helicobacter poly
39	115.5	7.8	675	22	ABH71143	Drosophila melanog
40	115	7.7	461	14	AAU39354	Epip protein. Sta
41	115	7.7	461	20	AAU43441	S. epidermis readi
42	115	7.7	889	22	ABP45918	S. enterica serova
43	115	7.7	1537	22	ABP58058	Drosophila melanog
44	114.5	7.7	674	22	AAU92775	Human protein sequ
45	114.5	7.7	2368	22	AAU34139	Staphylococcus aur

ALIGNMENTS

RESULT 1
AAV01365
ID AAV01365 standard; protein; 293 AA.
AC AAV01365;
XX
XX
DT 03-JUN-1999 (first entry)
XX
DE Wild-type Staphylococcal alpha haemolysin (HL) polypeptide.
XX
XX Staphylococcal; alpha-haemolysin; alphaHL; mutant; metal; biosensor;
KW heptameric pore assembly; micronutrient analysis; industrial effluent;
KW organic compound; explosive; macromolecule; bacteria; virus.
XX
OS Staphylococcus aureus.
XX
XX WO9905167-A1.
XX
PD 04-FEB-1999.
XX
PF 24-JUL-1998; 98WO-US15354.
XX
PR 25-JUL-1997; 97US-0053737.
XX
XX (UYMA-) UNIV MASSACHUSETTS.
XX
PI Bayley H, Braha O, Gouaux E, Kasianowicz J;
XX
DR WPI; 1999-153311/13.
XX
PT New mutant staphylococcal alpha-haemolysin - comprises a
PT heterologous amino acid that binds to analyte, particularly metal
PT ions

DR WPI; 2001-611495/70.
XX N-PSDB; AAS55423.
PT New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
PS Example 3; Seq ID No 13157; 51lpp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 327 AA;

Query Match 26.7%; Score 395.5; DB 22; Length 327;
Best Local Similarity 30.0%; Pred. No. 1.1e-22;
Matches 82; Conservative 70; Mismatches 112; Indels 9; Gaps 6;

QY 18 TTVKGTGLVTVYDKENGHHKVFYSFIDDKHNKKLLVIRTKGTIAGQYRVYSEBANKSG 77
DB 43 TLKTTATSDNDKUN-ISOILTFNFKDKSYDKDTLVKAAGNINSYGKPNPKDYNYSQ 101

QY 78 LAMPFAKVLQLPDNEVAQISDYPRNSIDTXEXMXTXGXNXXTXDXTKKXGLXG 137
DB 102 FYNGGKYNVSSESDAVNVVDYAPKN--QNEFVQOVLGYSG--GDINTSGLSGG 157

QY 138 XN--XSGXYKXVQDPFKTILSPDT-KKVGWKFVFNMMVNQMWGPDYDRDSWNPVYGNQ 194
DB 158 LNSKSFSEITNFKQESYRTTIDRKTNHKSTGCGVEAHKIMNNGWGPYGRDSDPTGYNE 217

QY 195 LEWKTRNGSKAADNFDLPNPKASSLSGSGSPDFATVITMDRASKQOQTNDIYIERVRD 254
DB 218 LFLGGROSSNAGQNFILPTHQMPLLARGNFNPEFISVLSHKQNDTK-KSKIKVTYQREMD 276

QY 255 DYQLHWTSTNWKGTNTKDKWTDSSERYKIDWE 287
DB 277 RYTNQWRLHWGNNYKNQVTFVTSYEDWQ 309

RESULT 7
AAU34287
ID AAU34287 standard; Protein; 325 AA.
XX
XX AAU34287;
XX
XX 14-FEB-2002 (first entry)
XX Staphylococcus aureus cellular proliferation protein #563.
DE Antisense; prokaryotic cellular proliferation protein;
XX antibiotic; antibacterial; drug design.
KW
KW Staphylococcus aureus.
OS
XX WO200170955-A2.
PN
XX 27-SEP-2001.

XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Haseibeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR N-PSDB; AAS52146.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 5783; 51lpp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 325 AA;

Query Match 26.1%; Score 388; DB 22; Length 325;
Best Local Similarity 29.9%; Pred. No. 4.4e-22;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;

QY 9 TGTDTI-GSNTTV--KTGDLVTVYDKENG-----MHKKVFYSFIDDKHNKKLLVIR 56
DB 21 SGTANAEGKITPVSVKVKVDKVTLYKTATADSKFKISQILTFNFKDKSYKDXDTLVLK 80

QY 57 TKGTIAGQYRVYSEBANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEXMXT- 115
DB 81 ATGNINSGFVKPNPDYDFSKLWYGAKNVYSISSQSDNSVNVVDYAPKNQNEEFQVQNTL 140

QY 116 -XTAGXNXXTXDXTKKXGLXGXNXXSGXTXKXVQDPFKTILSPDT-KKVGWKFVFN 173
DB 141 GYTFGGDISISNGL---SGGLNG-NTAFSETINIKQESYRTTILSRNTNRYNMGVGAHK 196

QY 174 MVNONNGPYDRDSWNPVYGNOLFMKFTNGSMKAADNFDLPNPKASSLSGSGSPDFAVIT 233
DB 197 IMNNGWGPYGRDSFHPPTYGNEFLAGQSSAYAGQNFIAHQHMPLLSRNFPNPEFLSVLS 256

QY 234 MDRKASKQOQTNDIYIERVRDDYQLHWTSTNWKGTNTKDKWTDSSERYKIDWEKEM 291
DB 257 -HRQDGAKKSKIIVTYQREMDLIQIRWNGFYWAGANYKNKFTKTRFKSYEIDWENHKV 313

RESULT 8
AAU37100


```
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 325 AA;

Query Match 26.1%; Score 388; DB 22; Length 325;
Best Local Similarity 29.9%; Pred. No. 4.4e-22;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;

QY 9 TGTDTI-GSNTTV---KTGDLVTYDKENG-----MHKKVYFSFIDDKNNHKKLLVIR 56
Db 21 SGTANAEKGTTPSVKVKVDKVLTKYTATADSDKFKISQILTFNFKDKSYDKDLVLK 80
QY 57 TKGTIAGQYRVYSEEGANKSGLWPSAFKVLQLPDNEVAQISDYPRNSIDTXEMXT- 115
Db 81 ATGNINSGFVKPNPDYDFSKLYGAKYVNSISSQNSDVNVVDYAPKQNEEFQVQNTL 140
QY 116 -XTXGXNXXXTDXTKXGXLCXNXXSGTXKXKXQVDPFKTILESPTD-KKVGWKVIFNN 173
Db 141 GYTFGGDISISNGL---SGGLNG-NTAFSETINYKQESYRTTILSRNTNYKNVGVGEAHK 196
QY 174 MVNQNWGPYDRDSNPNVYGNQLFPMKTRNGSMKAADNFDLPNKASSLLSGFSPDFATVIT 233
Db 197 IMNNGWGPYGRDSFHTYTGNEFLAGRSAYAGQNFIAHQHMPLLSRSNFEFLSVLS 256
QY 234 MDRKASKQQTINIDVIYERVDRDDYQLHWTSTNWKGTNTKDKWTDSSERYKIDWEKEEM 291
Db 257 -HRQDGAKSKITVTYQREMDLYQIRWNGFYWAGANYKFKTRTKSTYEDIDWENHKV 313

RESULT 11
AAU75478
ID AAU75478 standard; Protein; 325 AA.
XX
AC AAU75478;
XX
DT 23-APR-2002 (first entry)
XX
DE S. aureus antigenic protein #4.
XX
KW Antigenic protein; vaccine; SEREX; antibacterial;
KW antiinflammatory; dermatological; antiulcer; tuberculostatic;
KW immunosuppressive; septicaemia; food poisoning; skin disorders;
KW peritonitis; endocarditis; tuberculosis; blood infection; sepsis;
KW meningitis; pneumonia; stomach ulcer; gonorrhea; necrotising fasciitis;
KW impetigo; Lyme's disease; gastro-enteritis; dysentery; shigellosis.
XX
OS Staphylococcus aureus.
XX
PN WO200198499-A1.
XX
DT 27-DEC-2001.
XX
PF 20-JUN-2001; 2001WO-GB02685.
XX
PR 20-JUN-2000; 2000GB-0014907.
XX

CC The invention relates to antisense inhibitors of genes essential to
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:41:14 ; Search time 20.5 Seconds
(without alignments)
420.532 Million cell updates/sec

Title: CLAIM7

Perfect score: 1484

Sequence: 1 ADSDINIKTGTDIGSNTV.....WDRSSERYKIDWEKEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	293	19.7	336	US-08-666-405-28	Sequence 28, Appl
2	133.5	9.0	10182	US-09-134-001C-3159	Sequence 3159, Ap
3	124.5	8.4	1231	US-09-071-035-420	Sequence 420, App
4	124.5	8.4	1265	US-09-071-035-418	Sequence 418, App
5	118	8.0	447	US-09-627-376-10	Sequence 10, Appl
6	118	8.0	1584	US-09-457-040B-27	Sequence 27, Appl
7	115	7.7	461	US-08-392-625-24	Sequence 24, Appl
8	115	7.7	461	US-08-466-961A-24	Sequence 24, Appl
9	113.5	7.6	246	US-08-645-193B-26	Sequence 26, Appl
10	113.5	7.6	246	US-08-276-151-7	Sequence 7, Appl
11	112	7.5	512	US-08-856-253-6	Sequence 6, Appl
12	112	7.5	1183	US-08-447-031A-2	Sequence 2, Appl
13	110.5	7.4	1094	US-09-268-347-32	Sequence 32, Appl
14	107	7.2	2314	US-09-268-347-49	Sequence 49, Appl
15	106.5	7.2	1751	US-09-136-574A-44	Sequence 44, Appl
16	104.5	7.0	624	US-08-947-965-78	Sequence 78, Appl
17	104.5	7.0	655	US-08-469-202-28	Sequence 28, Appl
18	104.5	7.0	655	US-08-484-434C-35	Sequence 35, Appl
19	104.5	7.0	679	US-08-913-942-15	Sequence 15, Appl
20	104.5	7.0	679	US-09-268-347-26	Sequence 26, Appl
21	104	7.0	464	US-09-426-072-2	Sequence 2, Appl
22	103.5	7.0	1435	US-08-568-459A-4	Sequence 4, Appl
23	103.5	7.0	1435	US-08-487-8268-4	Sequence 4, Appl
24	103.5	7.0	1435	US-09-210-288-4	Sequence 4, Appl
25	103.5	7.0	1833	US-08-621-944A-4	Sequence 4, Appl
26	103.5	7.0	1833	US-08-945-567D-4	Sequence 4, Appl
27	103.5	7.0	1992	US-08-621-944A-3	Sequence 3, Appl

28 103.5 7.0 1992 4 US-08-945-567D-3 Sequence 3, Appli
29 103.5 7.0 2048 4 US-09-268-347-48 Sequence 48, Appl
30 102.5 6.9 345 4 US-08-856-253-7 Sequence 7, Appl
31 102.5 6.9 655 1 US-08-469-202-27 Sequence 27, Appl
32 102.5 6.9 933 3 US-08-484-434C-34 Sequence 34, Appl
33 102.5 6.9 933 3 US-08-293-728-2 Sequence 2, Appl
34 102.5 6.9 933 4 US-09-421-868-2 Sequence 2, Appl
35 102.5 6.9 1287 1 US-08-200-232-2 Sequence 2, Appl
36 102.5 6.9 1287 5 PCT-US95-02219-2 Sequence 2, Appl
37 102.5 6.9 1287 5 PCT-US95-02219A-2 Sequence 2, Appl
38 102.5 6.9 1401 4 US-09-127-670-6 Sequence 6, Appl
39 102 6.9 1098 1 US-08-409-995-2 Sequence 2, Appl
40 102 6.9 1098 3 US-08-685-467-2 Sequence 2, Appl
41 102 6.9 1098 4 US-09-377-155-32 Sequence 32, Appl
42 102 6.9 1098 4 US-08-913-942-2 Sequence 2, Appl
43 102 6.9 1098 4 US-09-669-974-32 Sequence 32, Appl
44 102 6.9 1098 4 US-09-268-347-44 Sequence 44, Appl
45 101.5 6.8 600 4 US-09-388-743-22 Sequence 22, Appl

ALIGNMENTS

RESULT 1

US-08-666-405-28
; Sequence 28, Application US/08666405
; Patent No 5874220

; GENERAL INFORMATION:

; APPLICANT: FACH, Patrick; GUILLOU, Michel

; TITLE OF INVENTION: PRIMERS FOR THE

; TITLE OF INVENTION: AMPLIFICATION OF GENES CODING FOR THE

; TITLE OF INVENTION: ENTEROTOXIN AND THE LEICITHINASE OF CLOSTRIDIUM

; TITLE OF INVENTION: PERRINGENS AND THEIR APPLICATION TO THE

; TITLE OF INVENTION: DETECTION AND NUMERATION OF THESE BACTERIAE

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIERMAN & MUSERLIAN

; STREET: 600 THIRD AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10016

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/666,405

; FILING DATE: 08-NOV-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP94/04292

; FILING DATE: 22-DEC-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/172,026

; FILING DATE: 22-DEC-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: MUSERLIAN, CHARLES A

; REGISTRATION NUMBER: 19,683

; REFERENCE/DOCKET NUMBER: 102.164

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 661-8000

; TELEFAX: (212) 661-8002

; INFORMATION FOR SEQ ID NO: 28:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 336 amino acids

; TYPE: amino acids

; STRANDEDNESS: single

; TOPOLOGY: linear

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

Db 601 ----YKGSILNLTMTTKAPSYQVYDDNDL 629

RESULT 4

US-09-071-035-418

Sequence 418, Application US/09071035

Patent No. 6448043

GENERAL INFORMATION:

APPLICANT: Gil H. Choi

TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 496

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/071,035

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: A. Anders Brookes

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB369P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 418:

SEQUENCE CHARACTERISTICS:

LENGTH: 1265 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-071-035-418

Query Match 8.4%; Score 124.5; DB 4; Length 1265;

Best Local Similarity 21.2%; Pred. No. 2.5;

Matches 58; Conservative 50; Mismatches 108; Indels 57; Gaps 11;

QY 40 YSFIDKNNKLLVIRTKGTAG-----QRYVYSEEGANKSGLAWPSAFK- 85

Db 427 YNLTNENKNTVMSSNLGTDTATLKNLSVPVNFQYRVNTFFYGASDTFTLPKRYKS 486

QY 86 VOLQLPDNEVAQISDYPRNSIDTXEXMXTXGXNKNXTDXTKKXGLXGXNKSXGXT 145

Db 487 INITKSGDGTDPAPLPKLYNIDQVEMHMPVITYN-----KLKQLSG--QTFGFN 535

QY 146 KXVQPDF--KTL--ESPDKKVGKVIFFNNMVNQWGPYDRDSWNPVYGNQLFMKTRN 201

Db 536 ALADQPEFTYKTLFGESGIDDPVNYTWS-----GPVY---YYLENR- 575

QY 202 GSKAADNFDLPNKASLLSSGSPDFATVITMDRASKAQQTNIIDVIYERVDYQLH-W 260

Db 576 ---KVTENFVDNTGAKITPTGTQCKKTVITSDATYTFKQAGTLPDTYTTGKTYFKRGW 632

QY 261 TSTNWKGTNTKDKWDRSSERYKIDWKEMTN 293

Db 633 ----YKGSILNLTMTTKAPSYQVYDDNDL 661

RESULT 5

US-09-627-376-10

Sequence 10, Application US/09627376

Patent No. 6342385

GENERAL INFORMATION:

APPLICANT: Qi, Fengxia

TITLE OF INVENTION: MUTACIN I BIOSYNTHESIS GENES AND PROTEINS

FILE REFERENCE: UAB-17402/22

CURRENT APPLICATION NUMBER: US/09/627,376

CURRENT FILING DATE: 2001-05-30

NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn version 3.0

SEQ ID NO 10

LENGTH: 447

TYPE: PRT

ORGANISM: Streptococcus mutans

US-09-627-376-10

Query Match 8.0%; Score 118; DB 4; Length 447;

Best Local Similarity 20.5%; Pred. No. 0.95;

Matches 68; Conservative 66; Mismatches 102; Indels 96; Gaps 19;

QY 2 DSDINIKTGTDTIGSNTTGYKGLVYDKENGMMHKVYFSDDKNNKKLL-VIRTKGT 60

Db 142 DVDHNAFTGMIDRSKNFVAGG---YDNSESSETGNINDIDDKGHTAVAGQTAANGQ 198

QY 61 IAGQ-----YRVYSEEGANKSGLAW-----PSAFKVQL---QLPDNEVA 96

Db 199 IFGVSPGTNLLIYRVF---GKSKECWILKALIDATNNGANVINLSLOYIKIPNGDIW 255

QY 97 QIS-----DYPRNSIDTXEXMXTXGXNKNXTDXTKKXGLXGXNKSXGXTK 147

Db 256 EGAELGYKFAIDYATRHNV--IVVAATGNDGLSDONGEVKTYNQSQSDMSQNDTVE 313

QY 148 XVQPDFKTTLESPDKKVG-----WKVIFNNMVNQ---NW-----GPYDRDSW 187

Db 314 ----DYPSSL--PNAIVAGSSDNNNQSRFSNYNQQDNFILAPGGGTTLLDOYGOEAW 367

QY 188 NPVYGNQLFMK-----TRNGSKAADNFDLPNKASLLSSGSPDFATVIT---MDRKA 238

Db 368 ---YNQKLFMEQVLTSTNNGVNDYAD-----GTSISTGKVGSGELAEIISNVHLQGDS 417

QY 239 SKQQTNIIDVIYERVDYQLHWTSTNWKGTNT 270

Db 418 SKARS---ILLN-----QVNYTSDGYKEIST 440

RESULT 6

US-09-457-040B-27

Sequence 27, Application US/09457040B

Patent No. 6387641

GENERAL INFORMATION:

APPLICANT: Vertex Pharmaceuticals Incorporated

APPLICANT: Bellon, Steve

TITLE OF INVENTION: Crystallized P38 Complexes

FILE REFERENCE: VPI/98-14

CURRENT APPLICATION NUMBER: US/09/457,040B

CURRENT FILING DATE: 1999-12-08

NUMBER OF SEQ ID NOS: 41

SOFTWARE: PatentIn version 3.0

SEQ ID NO 27

LENGTH: 1584

TYPE: PRT

ORGANISM: Dictyostelium Discoideum

US-09-457-040B-27

Query Match 8.0%; Score 118; DB 4; Length 1584;

Best Local Similarity 20.7%; Pred. No. 13;

Matches 36; Conservative 48; Mismatches 50; Indels 40; Gaps 6;

QY 93 NEVAQI-----SDYPRNSIDTXEXMXTXGXNKNXTDXTKKXGLXGXNKSXGXTK 147

Db 967 NRMIQIWNKSPDSYPKIAIDSSDKIRWPASGSGG---INISGWIGSSSSGSDGIT 1023

QY 148 XVQPDFKTI--LESPTDKKVGKVIFFNNMVNQWGPYDRDSWNPVYG-----NQLFMKTRN 201

```
Db 1024 EISSSSKNIRPKYSYQKEI-----EDNRRTISGGEKKNKYI----- 1063
QY 202 GSKMAADNPLDPNKASSLLSSGFSDFATVITMDRKASKQQTNDIVYRVRDD 255
Db 1064 -----DNQMDPHQIGSDGLLDFGQGGPPDEKNSKSTLSNEQIRYLOQRKD 1111

RESULT 7
US-08-392-625-24
; Sequence 24, Application US/08392625
; Patent No. 5837485
; GENERAL INFORMATION:
; APPLICANT: Entian, Karl-Dieter
; APPLICANT: G tz, Friedrich
; APPLICANT: Schnell, No. 5837485bert
; APPLICANT: Augustin, Johannes
; APPLICANT: Engelke, Gerhard
; APPLICANT: Rosenstein, Ralf
; APPLICANT: Kaletta, Cortina
; APPLICANT: Klein, Cora
; APPLICANT: Wieland, Bernd
; APPLICANT: Kupke, Thomas
; APPLICANT: Jung, G nther
; APPLICANT: Kellner, Roland
; TITLE OF INVENTION: Biosynthetic Process For The Preparation
; OF CHEMICAL COMPOUNDS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/392,625
; FILING DATE: 30-APR-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/876,791
; FILING DATE: 30-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0652.0980002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2540
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-392-625-24

Query Match 7.7%; Score 115; DB 2; Length 461;
Best Local Similarity 21.7%; Pred. No. 1.8;
Matches 63; Conservative 57; Mismatches 118; Indels 52; Gaps 13;

QY 4 DINKTGTDTGS-----NTTVKYGDLVTDYDKENGMMKKKFFYFIDDKHNHKKLLV--IR 56
Db 149 DSGVNSSHTDLKINKIVNPKNGFRGSENDESGN-----NFEEDKLNHGTLVAGQIG 203
QY 57 TKGTIAG-----QYRVYSEEGANKSGLAWPSAFKVLQQLPDNEVAQIS-----DYPR 104
Db 204 ANGLKGVNPGVEMNRYRVFGSKASEMLWLSKGIIDAANDNDNIVNLSGLNYLIKDNQNK 263
```

```
QY 105 NSIDTXEMXTXTGXNANXTDXDTXKXGLXGNXSGXTXKXVQP-DFKTILESPTDK 163
Db 264 KKLDRDEKVDYDALQKAINYAQ---KKSIVVAAGVNDGINVKVKKEINKRNLNSKTSK 320
QY 164 KVGWKVIFNNMNQN---WGPYDRDSWNPV---YGNOLF-MKTRNGSMKAADNPLDPNK 215
Db 321 KV-----YDSPANLNNVMTVGSIDNDYISEFSNYGNFNFDLMTTIGGSYK-----LLDKYG 371

QY 216 ASSLLSSGFSDFATVITMDRKASKQQTNDIVYRVR-----DDYQL 258
Db 372 KDAWLEKGYMQKQSVLSTSSNGRYIQSGTSLAAPKVS GALALEIDKYQL 421

RESULT 8
US-08-466-961A-24
; Sequence 24, Application US/08466961A
; Patent No. 5843709
; GENERAL INFORMATION:
; APPLICANT: Entian, Karl-Dieter
; APPLICANT: G tz, Friedrich
; APPLICANT: Schnell, No. 5843709bert
; APPLICANT: Augustin, Johannes
; APPLICANT: Engelke, Gerhard
; APPLICANT: Rosenstein, Ralf
; APPLICANT: Kaletta, Cortina
; APPLICANT: Klein, Cora
; APPLICANT: Wieland, Bernd
; APPLICANT: Kupke, Thomas
; APPLICANT: Jung, G nther
; APPLICANT: Kellner, Roland
; TITLE OF INVENTION: Biosynthetic Process for the Preparation of
; CHEMICAL COMPOUNDS
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,961A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/392,625
; FILING DATE: 22-FEB-1995
; APPLICATION NUMBER: US 07/876,791
; FILING DATE: 30-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/784,234
; FILING DATE: 31-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0652.0980004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-466-961A-24

Query Match 7.7%; Score 115; DB 2; Length 461;
```


Db	149	DSGVNSSHTDLKSINKIVNEVPKNFRGSEDSGNK-----NFEEDKLHGHTLVAGQIG	203
QY	57	TGKTIGAG-----QRYVTEEGANKSGIAWPSAFKVQLQLPDNPVAQIS-----DYYP	104
Db	204	ANGNLKGVNPGVENNVYRVFGSKSEMLWYSGKIIDAANDNDNDVINVSLGNYLTKNQNK	263
QY	105	NSIDTQEXMKATXTXGKXNNKTXDXTKKGLXGLXGNKXSGXTXKKXVQP-DEFTLIESPTDK	163
Db	264	KKLRDEKVDYDALOKAINYAQ--KKGSIVVAAGVNDGNVKKVKKEINKRNLSKTSK	320
QY	164	KVGHKVITFNKNVQN-----RGPYDRDSWNPV--YGNQLF-MKTRNGSKMAAADNFDLPNK	215
Db	321	KV-----YDSPANLNNVMTVGSIDDDNDYISEFSNYGNFIDLMTIGGSK-----LDDKYG	371
QY	216	ASSLLSSGFSFPDFATVITMDRKASKQOTNIDVIERVYR-----DDYQL	258
Db	372	KDAMLEKGYKQKSVLSTSSNGRYIIYOSGTSLAAPKVSYGALAEIDKYQII	421

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match	7.6%	Score 113.5;	DB 1;	Length 246;
	Best Local Similarity	19.1%;	Pred. No. 0.63;		
	Matches 50;	Conservative 52;	Mismatches 87;	Indels 73;	Gaps
QY	28 YDRENGMHKKVVF5FIDDKNNKKLLVIRPK7GIAGQ----	YRVYS-----	ESGANKSGL	78	
Db	21 YDDMAAAKAVTEQGHLSNEERLLLSVAYKNVVGARRSRVRSIS	TEQKTERNEKKQOM	80		
QY	79 AWP5AEKVLQQLPD--NEVAQISDIY-----	PRNSI-----	DTXEMXMTXTGXNX	122	
Db	81 GKREXKIEAELODICNDVLELQKYLIPNATOPESKVFYLLKMKGD	FVRLYSVAGSDNK	140		
QY	123 NXXTDXDTXKXGLXGNKNSXGXTXKXVQDPFKTILFESPTDKKVG	KWKFVFNMMYNQNWGPY	182		


```
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: 1997US001/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1751 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-136-574A-44

Query Match          7.2%; Score 106.5; DB 4; Length 1751;
Best Local Similarity 19.7%; Pred. No. 1.4e+02;
Matches 62; Conservative 52; Mismatches 102; Indels 99; Gaps 16;

QY 27 TYDKENGHHKVFYFIDDKNHKKLLVIRKGTIAGQYRVY-SEEGANK----- 75
Db 543 TYIEGG---KISGPYVMDKNRNIYVLVDFSGT-----KIYPGGEVEHKKQAQFKISVP 594

QY 76 SGLAW-----PSAFKVQLQDPNEVAQISDYYPNSI-----DTXE 111
Db 595 QGYPDPTNDPSYKGLTSOLEKNK--YIAAYDNNNLVWGLEPGAATSTPAPTSTPTPTPT 652

QY 112 XMXTXTXGXNNXNTXDXTKXGXLXG-----XNXSXGXTXKXVQDPDKTILESPTDKKVG 166
Db 653 PPTVTATPTPTPTPTGSPGTGSGVKVLYKNNETSASTGSIKVPFKIVNGSS----- 707

QY 167 WKVFNMMVNQWNGPYDRD-----SNPYYGNOLFWMKTRNGSKAADNFDLPNKASSL 219
Db 708 -SVDLSRVKIRYWTYVDGDKPQSAVCDWAI-----GASNVTFFVK----- 748

QY 220 LSSGFS-PDFATVITMDRKAKQQTNIQVIVERVRRDDYQLHWTSTNWKGTNKKKWTDRS 278
Db 749 LSSGVSGADYILEVGFSSGAGLQPGKDT-----GDIQVRFNKNDWSNYNQADDWS---- 799

QY 279 SERYKIDWEKEEMTN 293
Db 800 -----W-LQSMTN 806
```

Search completed: March 4, 2003, 10:46:03
Job time : 25.5 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:42:29 ; Search time 10.5 Seconds

(without alignments)

1176.748 Million cell updates/sec

Title: CLAIM7

Perfect score: 1484

Sequence: 1 ADSIDINIKTGTIGSNTTV.....WTDSSRYKIDWEKEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 188354 seqs, 42170167 residues

Total number of hits satisfying chosen parameters: 188354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/ECTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	421.5	28.4	330	10	US-09-815-242-5293
2	421.5	28.4	333	10	US-09-815-242-12121
3	395.5	26.7	327	10	US-09-815-242-5901
4	395.5	26.7	327	10	US-09-815-242-12729
5	395.5	26.7	327	10	US-09-815-242-13157
6	388	26.1	325	10	US-09-815-242-5783
7	388	26.1	325	10	US-09-815-242-12693
8	388	26.1	325	10	US-09-815-242-12779
9	388	26.1	325	10	US-09-815-242-13070
10	383	25.8	323	10	US-09-815-242-5891
11	383	25.8	338	10	US-09-815-242-13102
12	253	17.0	321	10	US-09-815-242-5771
13	253	17.0	321	10	US-09-815-242-12694
14	253	17.0	321	10	US-09-815-242-12721
15	240	16.2	312	10	US-09-815-242-5267
16	132.5	8.9	1300	10	US-09-815-242-4903
17	132.5	8.9	1300	10	US-09-815-242-10906
18	119	8.0	791	9	US-10-055-364-41
19	118	8.0	447	12	US-10-047-676A-10

Sequence 5816, Ap
Sequence 12967, A
Sequence 5456, Ap
Sequence 12686, A
Sequence 5635, Ap
Sequence 12389, A
Sequence 6, Appli
Sequence 45, Appli
Sequence 5898, Ap
Sequence 13137, A
Sequence 2, Appli
Sequence 381, App
Sequence 5471, App
Sequence 12544, A
Sequence 4, Appli
Sequence 7, Appli
Sequence 136, App
Sequence 11, Appli
Sequence 32, Appli
Sequence 73, Appli
Sequence 9, Appli
Sequence 8, Appli
Sequence 15, Appli
Sequence 109, App
Sequence 9, Appli
Sequence 257, App

ALIGNMENTS

RESULT 1

US-09-815-242-5293

; Sequence 5293, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR FILING DATE: 2000-03-21

; PRIOR FILING DATE: 2000-03-21

; PRIOR FILING DATE: 2000-05-23

; PRIOR FILING DATE: 2000-05-26

; PRIOR FILING DATE: 2000-10-23

; PRIOR FILING DATE: 2000-10-23

; PRIOR FILING DATE: 2000-11-27

; PRIOR FILING DATE: 2000-12-22

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5293

; LENGTH: 330

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

; US-09-815-242-5293

Query Match

Best Local Similarity

28.4%; Score 421.5; DB 10;

30.8%; Pred. No. 4.2e-22;

Length 330;

Db	91	SSGLEILYPNGYWSSTLURPWSYSVSIVQNDVDDNTNTKVDFAPKNQDETDRVKYTYGKT	150
Qy	121	NXNXTDXTKKXGXGXNKSXGTXKXVQPDFKXITLESPT-DKKVGWGWYFNFMVQNQW	179
Db	151	GGNFTINQGLSGNIT-QEYNYSFISYQPSYFTLDDKPTSNKAVAWKVEANVKNWGY	209
Qy	180	GPYDRS---WNPYVGQLFMKTRNGSKAAADPNKAKSSLLSGSFSPDFATVITMDR	236
Db	210	-DHTRDATDGTGKVGSEITFLTRNGNLWARDNFTPKMKPVTYSEGFNPEFLTVMSHDK	268
Qy	237	KASQQQNIDVIYERVRDDQLHWTSTN----WGKVTNKDKWTRDSSERYKIDWE	287
Db	269	K-GGHSLEFVYRYKRRVMDDEIRNRYNRYNIGWYSGKNVKNVAKESLAALYEIDWK	322

RESULT 3
US-09-815-242-5901
; Sequence 5901, Application US/09815242
; Patent No. US20020061569A1

```

? APPLICANT: Zyskind, Judith W.
? APPLICANT: Wall, Daniel
? APPLICANT: Trawick, John D.
? APPLICANT: Carr, Grant J.
? APPLICANT: Yamamoto, Robert T.
? APPLICANT: Xu, H. Howard
? TITLE OF INVENTION: Identification of Essential Genes in
? TITLE OF INVENTION: Prokaryotes
? FILE REFERENCE: ELITRA.011A
? CURRENT APPLICATION NUMBER: US/09/815,242
? CURRENT FILING DATE: 2001-03-21
? PRIOR APPLICATION NUMBER: 60/191,078
? PRIOR FILING DATE: 2000-03-21
? PRIOR APPLICATION NUMBER: 60/206,848
? PRIOR FILING DATE: 2000-05-23
? PRIOR APPLICATION NUMBER: 60/207,727
? PRIOR FILING DATE: 2000-05-26
? PRIOR APPLICATION NUMBER: 60/242,578
? PRIOR FILING DATE: 2000-10-23
? PRIOR APPLICATION NUMBER: 60/253,625
? PRIOR FILING DATE: 2000-11-27
? PRIOR APPLICATION NUMBER: 60/257,931
? PRIOR FILING DATE: 2000-12-22
? PRIOR APPLICATION NUMBER: 60/269,308
? PRIOR FILING DATE: 2001-02-16
? NUMBER OF SEQ ID NOS: 14110
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 5901
? LENGTH: 327
? TYPE: prt
? ORGANISM: Staphylococcus aureus
US-09-815-242-5901

```

QY	78	LAWPSAFKVQLQLPDNEVAQISDYIYPRNSIDTXEXMXTXTGXNXXNXTDXTXKXGXLG	137
Db	102	FYWGKKYNSVSSSESNDVNVDPAPKN--QNEEFGQQTGLGYSG--GDINISNGLSG	157
QY	138	XN--XSGXTXKXVQDPFKTILSPDT-KVGVKVFIFNNVNNQNWGPGYDRDSNPVYGNQ	194
Db	158	LNKSGKGFSEYINTKQSGSYRTIDRKNNHKSIGMGVEAHKIMNNGWGPGRDSYDPTYGNE	217
QY	195	LFMKTRNGSKMAADNFDLPNKASSLSSGFSPOFATVITMDRKASKQOQTNIDVYVERVD	254


```
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5783
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5783

Query Match 26.1%; Score 388; DB 10; Length 325;
Best Local Similarity 29.9%; Pred. No. 9.8e-20;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;

Qy 9 TGTDDI-GSNTTV---KTGDLVYDKENG-----MHKKVYFSDIDKHNKLLVIR 56
Db 21 SGTANAEGKITPVSKYKVDKVTLYKTTATADSKFKISQILTFNFKDKSYDKDLVLK 80
Qy 57 TKGTTAGYRVYSEEGANKSLWPSAFKVLQLPDNEVAOISDYPRNSIDTXXEMXT- 115
Db 81 ATGNINSGFVPNPNDYDFSKLYGAKYVNSISSQSDNSVNVVDYAPKNQNEEFQVQNTL 140
Qy 116 -XTGXNXXNTXDXTKXGXGLXGNXSGXTXKXVQPDFKTILESPTD-KKVGKVI FNN 173
Db 141 GYTFGGDISISNGL---SGGLNG-NTAFSEIYNYKQESYRTLSRNTYKNVGVGEA HK 196
Qy 174 MVNQWGPYDRDSNPVYGNQLFMKTRNGSKAADNFDLPNKASSLLSGSFSPDFATVIT 233
Db 197 IMNNGWGPYGRDSFHTYGNELFLAGROSSAYAGQNFIAQHMPLLSRNFPFLSVLS 256
Qy 234 MDRKASKQOTNIDVIYERVRDDYQLHWTSTNWKGTNTKDKWTDRSSERYKIDWEKEEM 291
Db 257 -HRQDGAKSKITVTYQREMDLYQIRWNGFYWAGANYKNFKTRTFKSYEIDWENHKV 313

RESULT 7
US-09-815-242-12693
; Sequence 12693, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; CURRENT APPLICATION NUMBER: US/09/815,242
```

```
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12693
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12693

Query Match 26.1%; Score 388; DB 10; Length 325;
Best Local Similarity 29.9%; Pred. No. 9.8e-20;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;

Qy 9 TGTDDI-GSNTTV---KTGDLVYDKENG-----MHKKVYFSDIDKHNKLLVIR 56
Db 21 SGTANAEGKITPVSKYKVDKVTLYKTTATADSKFKISQILTFNFKDKSYDKDLVLK 80
Qy 57 TKGTTAGYRVYSEEGANKSLWPSAFKVLQLPDNEVAOISDYPRNSIDTXXEMXT- 115
Db 81 ATGNINSGFVPNPNDYDFSKLYGAKYVNSISSQSDNSVNVVDYAPKNQNEEFQVQNTL 140
Qy 116 -XTGXNXXNTXDXTKXGXGLXGNXSGXTXKXVQPDFKTILESPTD-KKVGKVI FNN 173
Db 141 GYTFGGDISISNGL---SGGLNG-NTAFSEIYNYKQESYRTLSRNTYKNVGVGEA HK 196
Qy 174 MVNQWGPYDRDSNPVYGNQLFMKTRNGSKAADNFDLPNKASSLLSGSFSPDFATVIT 233
Db 197 IMNNGWGPYGRDSFHTYGNELFLAGROSSAYAGQNFIAQHMPLLSRNFPFLSVLS 256
Qy 234 MDRKASKQOTNIDVIYERVRDDYQLHWTSTNWKGTNTKDKWTDRSSERYKIDWEKEEM 291
Db 257 -HRQDGAKSKITVTYQREMDLYQIRWNGFYWAGANYKNFKTRTFKSYEIDWENHKV 313

RESULT 8
US-09-815-242-12779
; Sequence 12779, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; CURRENT APPLICATION NUMBER: 60/242,578
```


;
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12779
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12779

Query Match 26.1%; Score 388; DB 10; Length 325;
Best Local Similarity 29.9%; Pred. No. 9.8e-20;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;

QY 9 TGTDTI-GSNTTV---KTGDLVTYDKENG-----MHKKVYFSFIDDKNNHKKLLVIR 56
Db 21 SGTANAEGKITPVSVKVKVDKVTLYKTATADSKFKISQLITFNFKDKSYDKDTLVLK 80

QY 57 TGTIAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTXEXMXT- 115
Db 81 ATGINSFGVKPNPDYDFSKLYGAKYNYVSISSQSNDSNVNVDYAPKNQNEEFQVQNTL 140

QY 116 -XTGXNKNXTDXTKXKXGLXGNXSGXTXKXVQPDFKTLTLESPTD-KKVGKVIENN 173
Db 141 GYTFGGDISISNGL---SGGLNG-NTAFSETINYKQESYRTLSRNTNYKNVGVGAHK 196

QY 174 MVNQNWGPYDRDSNPNVYGNQLFMKTRNGSKMAADNFDLPNKASSLLSGFSPDFATVIT 233
Db 197 IMNNGWGPYGRDSFHPYTYGNELFLAGQSSAYAGQNFIAQHOMPLLSRSNFNPEFLSVLS 256

QY 234 MDRKASQOQNIDVIYERVDYQLHWTSTNWKGTNTKDKWTDSSRYKIDWEKEEM 291
Db 257 -HRQDGAKSKITVTYQREMDLYQIRWNGFYWAGANYKNEKTRTFKSTYEIDWENHKV 313

RESULT 9
US-09-815-242-13070
; Sequence 13070, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110

;
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13070
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13070

Query Match 26.1%; Score 388; DB 10; Length 325;
Best Local Similarity 29.9%; Pred. No. 9.8e-20;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;

QY 9 TGTDTI-GSNTTV---KTGDLVTYDKENG-----MHKKVYFSFIDDKNNHKKLLVIR 56
Db 21 SGTANAEGKITPVSVKVKVDKVTLYKTATADSKFKISQLITFNFKDKSYDKDTLVLK 80

QY 57 TGTIAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTXEXMXT- 115
Db 81 ATGINSFGVKPNPDYDFSKLYGAKYNYVSISSQSNDSNVNVDYAPKNQNEEFQVQNTL 140

QY 116 -XTGXNKNXTDXTKXKXGLXGNXSGXTXKXVQPDFKTLTLESPTD-KKVGKVIENN 173
Db 141 GYTFGGDISISNGL---SGGLNG-NTAFSETINYKQESYRTLSRNTNYKNVGVGAHK 196

QY 174 MVNQNWGPYDRDSNPNVYGNQLFMKTRNGSKMAADNFDLPNKASSLLSGFSPDFATVIT 233
Db 197 IMNNGWGPYGRDSFHPYTYGNELFLAGQSSAYAGQNFIAQHOMPLLSRSNFNPEFLSVLS 256

QY 234 MDRKASQOQNIDVIYERVDYQLHWTSTNWKGTNTKDKWTDSSRYKIDWEKEEM 291
Db 257 -HRQDGAKSKITVTYQREMDLYQIRWNGFYWAGANYKNEKTRTFKSTYEIDWENHKV 313

RESULT 10
US-09-815-242-5891
; Sequence 5891, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5891
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5891

Query Match 25.8%; Score 383; DB 10; Length 323;

```
Best Local Similarity 28.6%; Pred. No. 2.2e-19;
Matches 82; Conservative 79; Mismatches 108; Indels 18; Gaps 8;

QY 13 DIGSNNTVKTGLVTDKENGHHKVFYSFIDDKNNHKKLLVIRTKGTIAGQYRVYSEEG 72
Db 36 NLDGDTKMYTRATTSDSQKNITQSQFNFLTEPNYDKETVFIKAGTTIGSLRLDPNG 95

QY 73 ANKSLGAWPSAFKVLQ-LPDNEVAQISDYPRNSIDTXEMXMTXTXGXNXXNTXDXTXK 131
Db 96 YWNSTLRWPGSYSVSQNVDDNNNTNVTDFAPKNODESREVYK----YGYKTGGDFSN 151

QY 132 XGXLG---XNXSGXTXKXVQDPDKTILESTPDKK-VGKVV---IFNNMVNQNNGPYDR 184
Db 152 RGCLTGNITKESNYSETISYQOPSYRTLLDQSTSHKGVGKVEAHLINNNHGHDTQLTN 211

QY 185 DSWNPVYGQLPMKTRNGSMKAADNFPDNKASSLLSSGFSDFATVITMDRKASKQOQTN 244
Db 212 DSDNRT-KSEISLFRNGNLWAKDNFTPKDKMPVTVSEGFNFEFLAVMWSHDKK-DGKRSQ 269

QY 245 IDVIYERDDYQLHWTSTN----WKGNTNTKDKWTDNRSSERYKIDWE 287
Db 270 FVHYKRSMDERFKIDNRRHGFNGWSNGENHVDKKEEKLALYEVDMK 316

RESULT 11
US-09-815-242-13102
; Sequence 13102, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13102
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13102

Query Match 25.8%; Score 383; DB 10; Length 338;
Best Local Similarity 28.6%; Pred. No. 2.4e-19;
Matches 82; Conservative 79; Mismatches 108; Indels 18; Gaps 8;

QY 13 DIGSNNTVKTGLVTDKENGHHKVFYSFIDDKNNHKKLLVIRTKGTIAGQYRVYSEEG 72
Db 42 NLDGDTKMYTRATTSDSQKNITQSQFNFLTEPNYDKETVFIKAGTTIGSLRLDPNG 101

QY 73 ANKSLGAWPSAFKVLQ-LPDNEVAQISDYPRNSIDTXEMXMTXTXGXNXXNTXDXTXK 131
Db 96 YWNSTLRWPGSYSVSQNVDDNNNTNVTDFAPKNODESREVYK----YGYKTGGDFSN 151

QY 132 XGXLG---XNXSGXTXKXVQDPDKTILESTPDKK-VGKVV---IFNNMVNQNNGPYDR 184
Db 152 RGCLTGNITKESNYSETISYQOPSYRTLLDQSTSHKGVGKVEAHLINNNHGHDTQLTN 211

QY 185 DSWNPVYGQLPMKTRNGSMKAADNFPDNKASSLLSSGFSDFATVITMDRKASKQOQTN 244
Db 212 DSDNRT-KSEISLFRNGNLWAKDNFTPKDKMPVTVSEGFNFEFLAVMWSHDKK-DGKRSQ 269

QY 245 IDVIYERDDYQLHWTSTN----WKGNTNTKDKWTDNRSSERYKIDWE 287
Db 270 FVHYKRSMDERFKIDNRRHGFNGWSNGENHVDKKEEKLALYEVDMK 316

RESULT 12
US-09-815-242-5771
; Sequence 5771, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5771
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5771

Query Match 17.0%; Score 253; DB 10; Length 321;
Best Local Similarity 24.9%; Pred. No. 3.6e-10;
Matches 71; Conservative 72; Mismatches 118; Indels 24; Gaps 12;

QY 13 DIGSNNTV--KTGDLVTDKENGHHKVFYSFIDDKNNHKKLLVIRTKGTIAGQYRVYSE 70
Db 47 DIGQGAETIKRTQDITS--KRLAITQNIQDFEVKDKKYNKDALVVKMOGFSSR-TTYS 103

QY 71 --EGANKSLGAWPSAFKVLQ-LPDNEVAQISDYPRNSIDTXEMXMTXTXGXNXXNTXD 128
Db 104 LKYPYIKRMWIPQYINISLTKDSNV-DLINYLPKNKIDSAD--VSOKLGYNIGNGFQS 160

QY 129 TXKXGXLGXNKSXCXTXKXVQDPDKTILESTPDKKGVKVFIFNNMVNQNNGPYDRDSN 188
Db 161 APSIGSGSFSNYS--KTISYNQKNYVTEVSQNSKGVKGVKANSFVTPN-----GOV 211

QY 189 PVYGNQLPMKTRNGSMKAADNFPDNKASSLLSSGFSDFATVITMDR-KASKQOQTNIDV 247
```

*er.

APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011a
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5267
LENGTH: 312
TYPE: PRT
ORGANISM: Staphylococcus aureus
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(312)
OTHER INFORMATION: Xaa - Any Amino Acid
US-09-815-242-5267

Query Match 16.2%; Score 240; DB 10; Length 312;
Best Local Similarity 23.7%; Pred. No. 2.9e-09;
Matches 71; Conservative 71; Mismatches 119; Indels 38; Gaps 12;
QY 9 TGTDTIGSNITV--KTGDLVYDKEGHHKVFYFSDIDKHNKKLLVIRTKGTAGQVR 66
DB 30 TWIENIGDGAIVKRTEDVSS--KKWGTQNVQDFVKKDKYKDALIVKMGQFINSRTS 87
QY 67 VYSEEGAN---KSGLAWSAFKVLQLPDNEVAQISDYPRNSIDTXEMXTTXGXNKN 123
DB 88 FSDVKGSGVELTKRWLPFYQNIGLTKDPNVSLI-NYLPKNKIETD--VQQTLCYNI 144
QY 124 XYDXTXKXGLXGXNXXGXGTXKXVQDPFKTILESPDTDKKVGKVIFFNNMNVQNWGPYD 183
DB 145 GNFSAPSTGGNGSFNYS--KTIISYTKSYSEVDKQNSKVKGVCANEFVTP----- 196
QY 184 RDSWNPVGNLQFMKTRNG-SMKAADNFDLPNKASLLSSGFSDFATVITMDRKASK-- 240
DB 197 -DCKASADRYLFVQSPNGPSGAREYFAPDNQLPPLVQSGFNPSFITLSHE-KGSXXI 254
QY 241 -----QQTNDIVYERVDYQLHWTSTNWKGTNTKDKWTDTS-SERYKIDWEKEM 291
DB 255 RYNLAFHGRNVDTIYATL-----FPRTGIYAERKHNAFVNRNFVYEVNWKTHEI 306

Search completed: March 4, 2003, 10:46:24
Job time : 11.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:40:49 ; Search time 27.5 seconds
(without alignments)
1024.269 Million cell updates/sec

Title: CLAIM7
Perfect score: 1484
Sequence: 1 ADSINIKTGTDTIGSNTTV.....WDRSSRYKIDWEKEEMTN 293

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1484	100.0	319	2 S69209	alpha-toxin precursor
2	1480	99.7	319	2 F89887	Alpha-Hemolysin pr
3	397	26.8	325	2 T00161	leucocidin chain F
4	394.5	26.6	327	2 B89968	leukotoxin, LukD [
5	394	26.5	325	2 S32212	leucocidin chain F
6	393	26.5	325	2 B49234	leucocidin R, comp
7	388	26.1	300	2 B49238	gamma-hemolysin ga
8	388	26.1	325	2 F90043	gamma-hemolysin co
9	385	25.9	338	2 C89991	hypothetical prote
10	384.5	25.9	326	2 S68224	synergohymenotropi
11	374	25.2	322	2 JC5469	Panton-Valentine l
12	374	25.2	323	2 J01530	leucocidin chain F
13	374	25.2	323	2 JN0627	leucocidin chain F
14	293	19.7	336	2 I40826	beta-toxin - Clost
15	262	17.7	308	2 JC5468	leucocidin chain l
16	254.5	17.1	351	2 D89991	hypothetical prote
17	253	17.0	309	2 D90043	gamma-hemolysin ch
18	252	17.0	309	2 JC4282	gamma-hemolysin II
19	248	16.7	286	2 C49238	gamma-hemolysin co
20	248	16.7	315	2 PC4078	hlgc-like protein
21	248	16.7	315	2 E90043	gamma-hemolysin co
22	247.5	16.7	311	2 C89968	leukotoxin LukE [i
23	246	16.6	321	2 S49271	hlga-like protein
24	245	16.5	315	2 A49234	leucocidin R S com
25	242	16.3	315	2 JN0626	leucocidin chain S
26	237	16.0	312	2 T00160	leucocidin chain S
27	233	15.7	312	2 S32211	leucocidin chain S
28	203.5	13.7	310	2 S68225	synergohymenotropi
29	129	8.7	1272	2 C90593	hypothetical prote

30	124	8.4	410	2 T13531	hypothetical prote
31	119	8.0	807	2 T42924	glycoprotein B - a
32	118	8.0	1243	2 S60138	sex factor aggrega
33	118	8.0	1584	2 T18276	protein-tyrosine k
34	117.5	7.9	1787	2 T20160	hypothetical prote
35	116.5	7.9	245	2 S13467	14-3-3 protein - b
36	116.5	7.9	1965	2 S75200	fat protein - Syne
37	116	7.8	435	2 G86907	D-alanyl-D-alanine
38	116	7.8	859	2 B64430	DNA-directed RNA p
39	116	7.8	868	2 E89897	conserved hypothe
40	116	7.8	2269	2 T28677	rhostry protein -
41	116	7.8	2893	2 A64556	toxin-like outer m
42	116	7.8	3083	2 AH2493	hypothetical prote
43	115.5	7.8	4936	2 AH2515	hypothetical prote
44	115	7.7	461	2 S23420	probable subtilisi
45	115	7.7	895	2 AD0541	outer membrane fil

ALIGNMENTS

RESULT 1

S69209
alpha-toxin precursor - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 06-Sep-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-2000
C:Accession: S69209
R:Hedengrann, G.
submitted to the EMBL Data Library, October 1992
A:Reference number: S69209
A:Accession: S69209
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <HED>
A:Cross-references: EMBL:X01645; NID:g46763; PIDN:CAA25801.1; PID:g46765
C:Superfamily: leukocidin
C:Keywords: toxin
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-319/Product: alpha-toxin #status predicted <MAT>

Query Match 100.0%; Score 1484; DB 2; Length 319;
Best Local Similarity 93.2%; Pred. No. 5.8e-84;
Matches 273; Conservative 20; Mismatches 0; Indels 0; Gaps 0;

QY	1	ADSDINIKTGTDTIGSNTTVKTKGDLVYDKENGMMHKVFYSFIDDKNNKKLLVIRTKGT	60
DB	27	ADSDINIKTGTDTIGSNTTVKTKGDLVYDKENGMMHKVFYSFIDDKNNKKLLVIRTKGT	86
QY	61	IAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNSIDTXEXMXTXTXGX	120
DB	87	IAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNSIDTXEXMXTXTXGX	146
QY	121	NXNXTXDTXXKXGXKXSGXTKXKXVOPDFKTIESTPTDKKVGKVFNNMNNQWG	180
DB	147	NGNVTGDDTKIGGLIGANVSIGHTLKTVQDFKTIESTPTDKKVGKVFNNMNNQWG	206
QY	181	PYDRSDSNVYGNQLPFMKTRNGSKAADNFDNPKASSLSGSPDFAVITMDRKASK	240
DB	207	PYDRSDSNVYGNQLPFMKTRNGSKAADNFDNPKASSLSGSPDFAVITMDRKASK	266
QY	241	QQTNIDVIYERVDYQLHWTSTNWKTNTKDKWDRSSERYKIDWEKEEMTN	293
DB	267	QQTNIDVIYERVDYQLHWTSTNWKTNTKDKWDRSSERYKIDWEKEEMTN	319

RESULT 2

F89887
Alpha-Hemolysin precursor [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 01-Mar-2002
C:Accession: F89887
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, T.

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F90043
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <KUR>
A:Cross-references: GB:BA000018; PID:g13702370; PIDN:BA843511.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: hlgB
C:Superfamily: leukocidin

Query Match 26.1%; Score 388; DB 2; Length 325;
Best Local Similarity 29.9%; Pred. No. 2.8e-16;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;
QY 9 TGTDTI-GSNTTV---KTGDLVTYKENG-----MHKKVFYSFIDDKNNHKKLLVIR 56
Db 21 SGTANAEGKTPVSVKVDDKVTLYKTATADSDKFKISQILTFNFKDKSYDKDFLVLK 80
QY 57 TKTIAQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNSIDTXEMXT- 115
Db 81 ATGNINSGEVFKPNPDYDFSKLYGAKYNVSISSQSDNSVNVVYAPKQNEEFQVQNTL 140
QY 116 -XTXGXNXXNTXDTXKXGLXGXNXXGXTXKXVQPDFKTTILESPD-KKVGWKFYFN 173
Db 141 GYTFGGDISINGL---SGGLNG-NFAFSETINIKQESTRTLSRNTNTKKNVGMGVEAHK 196
QY 174 MVNQNGPYDRDSWNPVYGNQLFMTKTRNGSMKAADNFLDPNKASSLLSGFSPDFATVIT 233
Db 197 IMNNGWPGYGRDSFHPFTYGNELFLAGRSSAVAGQNFIAHQHQLPLLSRNFENPEFLSVLS 256
QY 234 MDRKASKQQTNDIVYERVDDYQLHWTSTNWKGTNTKDKWTDSSRYKIDWEKEEM 291
Db 257 -HRQDGAKKSKITVTYQREMDLYQIRWNGFYWAGANYKNFKTRTFKSTYEIDWENHKV 313

RESULT 9
C89991
hypothetical protein SA1812 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89991
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C89991
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-338 <KUR>
A:Cross-references: GB:BA000018; PID:g13701799; PIDN:BA843092.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA1812
C:Superfamily: leukocidin

Query Match 25.9%; Score 385; DB 2; Length 338;
Best Local Similarity 28.2%; Pred. No. 4.7e-16;
Matches 82; Conservative 81; Mismatches 110; Indels 18; Gaps 8;
QY 13 DIGSNTVTKTGLVYDKENGMMHKVFYSFIDDKNNHKKLLVIRTKGTIAGQYRVYSEEG 72
Db 42 NLDGDTKMYTRTATTSDSKNTQSLQFNFLTEPNYDRETVFIRAKGTIGSGLRILDENG 101
QY 73 ANKSGLAWPSAFKVOLO-LPDNEVAQISDYYPNSIDTXEMXTYTXGNXXNTYDXTYK 131
Db 102 YNSTLRWPGSVSVSIQNVDDNNNTVTDFAPKNODESREVKYT----YGYTKTGDFDSIN 157

QY 132 XGXLXG---XNXXSGXTXKXVQPDFKTTILESPDKK-VGMKV---IFNNMVNQNWGPYDR 184
Db 158 RGLGNTIITKESNYSETISYQPSRYTLDDQSTSHKGVKGWKEAHLINNMGHDHTRQLTN 217
QY 185 DSWNPVYGNQLFMKTRNGSMKAADNFLDPNKASSLLSGFSPDFATVITMDRKASKQQTN 244
Db 218 DSDNRT-KSEIFSLTRNGNLWAKDNFTPKDKMPVTVYSEGFNPEFLAVMSHDKK-DKGKQ 275
QY 245 IDVIYERVDDYQLHWTSTN---WKGTNTKDKWTDSSRYKIDWEKEEM 291
Db 276 FVHYKRSMDKIDNRHGHFWYSGENHVDKKEKLSALYEDVWKTHDV 326

RESULT 10
S88224
synergohymenotropic toxin component F precursor - Staphylococcus intermedius
C:Species: Staphylococcus intermedius
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-2000
C:Accession: S68224; S78213; S44945
R:Prevost, G.; Bouakham, T.; Piemont, Y.; Montell, H.
FEBS Lett. 376, 135-140, 1995
A:Title: Characterisation of a synergohymenotropic toxin produced by Staphylococcus
A:Reference number: S68224; MUID:96105366; PMID:7498527
A:Accession: S68224
A:Molecule type: DNA
A:Residues: 1-326 <PRE>
A:Cross-references: EMBL:X79188; NID:9487685; PIDN:CAA55783.1; PID:9487687
A:Experimental source: ATCC 51874
A:Note: the authors translated the codon TAT for residue 103 as Ile
A:Accession: S78213
A:Molecule type: protein
A:Residues: 27-43 <PRE>
A:Experimental source: ATCC 51874
C:Genetics:
A:Gene: lukF-I
C:Superfamily: leukocidin
C:Keywords: toxin
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-326/Product: synergohymenotropic toxin component F #status experimental <MAT>

Query Match 25.9%; Score 384.5; DB 2; Length 326;
Best Local Similarity 30.1%; Pred. No. 4.7e-16;
Matches 83; Conservative 72; Mismatches 108; Indels 13; Gaps 9;
QY 18 TTVKGTGLVTDYKEN-GMHKKVFYSFIDDKNNHKKLLVIRTKGTIAGQYRVYSEEGANKS 76
Db 43 TLYKT--TATADSDNINISQLTFFNFKDKSYDKDTLVKAGNINSGTSKSPNPNDIYS 100
QY 77 GLAWPSAFKVLQLPDNEVAQISDYYPNSIDTXEMXTXGXNXXNTXGXNXXNTXGXNXX 136
Db 101 SFYWGAKYNSISAESKCAVNVVDYAPKN--QNEEFQVQNTLGYSEF--GDISISKGLSG 156
QY 137 GAN--XSXGTXKXVQPDFKTTILESPDKK-VGMKVIFNNMVNQNWGPYDRSWNPVYGN 193
Db 157 GLNGSEFSETINIKQESTRTIDKTDNKTIGWVEAHKIMNAGWPGYGRDSFHDLYGN 216
QY 194 QLFMKTRNGSMKAADNFLDPNKASSLLSG-FSPDFATVITMDRKASKQQTNDIVYERV 252
Db 217 ELFLGGROSKINAGQNF-L-PTROMPLLAGNFPNPEFLSVLSHKPNGAK-TSKIRVITYORE 274
QY 253 RDDYQLHWTSTNWKGTNTKDKWTDSSRYKIDWEK 288
Db 275 MDEYTNWNGFHWGCTNYKNQNNATFTSFYEIDWDQ 310

RESULT 11
JC5469
Panton-Valentine leukocidin LukF-pv chain precursor - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 20-Jun-2000
C:Accession: JC5469; PC4337
R:Kaneko, J.; Muramoto, K.; Kamio, Y.

A; Residues: 27-85 <KAM>
A; Experimental source: strain No.4 (RIMD 310925)
C; Comment: This leukocidin acts synergistically to induce cytotoxic changes in human C; Genetics:
A; Gene: lukF
C; Complex: heterodimer; chains S and F
C; Superfamily: leukocidin
C; Keywords: hemolysin, toxin
F; 1-25/Domain: signal sequence #status predicted <SIG>
F; 26-323/Product: leukocidin chain F #status experimental <LEU>

Query Match 25.2%; Score 374; DB 2; Length 323;
Best Local Similarity 29.6%; Pred. No. 2.1e-15;
Matches 88; Conservative 72; Mismatches 117; Indels 20; Gaps 8;

QY 9 TGTTDI-GSNTTV---KTGDLVTVDKENG-----MHKKVYFSFIDDKHNKKLLVTR 56
DB 21 SGTANAGKITPVSVKVKVDDKVTLYKTTATADSDRFKISQILTFNFIKDKSDYDKDFTLVLK 80
QY 57 TKGPTIAGQYRVYSEEGANKSLGAWPSAFKVLQLPDNEVAQISDYPRNSIDTXXEMXT- 115
DB 81 ATGNINSGFVKPNPDYDFSKLYMGAKYNSISSQSDNSVNAVDPYAPKNONEEFQVONTL 140
QY 116 -XTYGXNKNXTDYTXKXGLGXNXXSGXTYKXVQPFKTTILEPTDKKVGWVKVFNWM 174
DB 141 GYTFGGDISISNGL---SGGLNG-NTAFSETINYYQESYRTLSRNTNYKNVGWGEVAHKI 196
QY 175 VNQWNGPYDRDSWNPVYGNQLFMKTRNGSKMAADNFDLPNPKASSLLSSGFSQFDFATVITM 234
DB 197 MN-CWGPYGRDSFPHTYGNELFLAGRSAYAGQNFIAHQHMPLLSRNFNPRFSLVLSH 255
QY 235 DRKASKOQTNDIVYERVVDDYQLHWTSTNMKGTTKDKWTDRSERYKIDWEKEM 291
DB 256 RQDAAK-KSKITVTVQREMDLYQIRWNGFYWAGANYKNFKTRTEKSTYEIDWENHKV 311

RESULT 13
JN0627
leukocidin chain F precursor - Staphylococcus aureus
C; Species: Staphylococcus aureus
C; Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-2000
C; Accession: JN0627; S32416
R; Rahman, A.; Izaki, K.; Kamio, Y.
A; Title: Gamma-hemolysin genes in the same family with LukF and LukS genes in methicillin-resistant strains of Staphylococcus aureus
A; Reference number: JN0625; MUID:93372503; PMID:7763998
A; Accession: JN0627
A; Molecule type: DNA
A; Residues: 1-323 <RAH>
A; Cross-references: GB:S65052; NID:g410004; PIDN:AAC60446.1; PID:g410007
R; Kamio, Y.; Rahman, A.; Nariya, H.; Ozawa, T.; Izaki, K.
FEBS Lett. 321, 15-18, 1993
A; Title: The two staphylococcal bi-component toxins, leukocidin and gamma-hemolysin.
A; Reference number: S32416; MUID:93223836; PMID:8467905
A; Accession: S32416
A; Status: preliminary
A; Molecule type: protein
A; Residues: 27-85 <KAM>
C; Genetics:
A; Gene: lukF
C; Superfamily: leukocidin
C; Keywords: toxin
F; 1-26/Domain: signal sequence #status predicted <SIG>
F; 27-323/Product: leukocidin chain F #status predicted <MAT>

Query Match 25.2%; Score 374; DB 2; Length 323;
Best Local Similarity 29.3%; Pred. No. 2.1e-15;
Matches 87; Conservative 73; Mismatches 117; Indels 20; Gaps 8;

QY 9 TGTTDI-GSNTTV---KTGDLVTVDKENG-----MHKKVYFSFIDDKHNKKLLVTR 56
DB 21 SGTANAGKITPVSVKVKVDDKVTLYKTTATADSDRFKISQILTFNFIKDKSDYDKDFTLVLK 80

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:39:33 ; Search time 17 Seconds
(without alignments)
714.857 Million cell updates/sec

Title: CLAIM7

Perfect score: 1484

Sequence: 1 ADSIDINIKTGTDTIGSNTTV.....WDRSSERYKIDWEKEEMTN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1484	100.0	319	1 HLA_STAAU	P09616 staphylococ
2	388	26.1	325	1 HLGE_STAAU	Q07226 staphylococ
3	374	25.2	323	1 LURF_STAAU	P31715 staphylococ
4	253	17.0	309	1 HLGA_STAAU	P31714 staphylococ
5	248	16.7	315	1 HLGC_STAAU	Q07227 staphylococ
6	242	16.3	315	1 LUKS_STAAU	P31716 staphylococ
7	118	8.0	1584	1 KYK1_DICDI	P18160 dictyosteli
8	117.5	7.9	1787	1 CHD3_CABEL	Q22516 caenorhabdi
9	116.5	7.9	245	1 L43B_BOVIN	P29358 bos taurus
10	116	7.8	859	1 RPA2_METJA	Q58446 methanococc
11	115	7.7	461	1 EPIP_STAEP	P30199 staphylococ
12	114	7.7	464	1 MRJ4_APIME	Q17061 apis mellif
13	113.5	7.6	245	1 L43B_HUMAN	P31946 homo sapien
14	112	7.5	1183	1 CNA_STAAU	Q53654 staphylococ
15	111.5	7.5	1160	1 TFC3_YEAST	P34111 saccharomyc
16	111.5	7.5	1288	1 VACA_HELPJ	Q92kw5 helicobacte
17	110.5	7.4	245	1 L43B_MOUSE	Q9cqv8 mus musculu
18	109.5	7.4	798	1 PBPA_NEILA	Q87579 neisseria l
19	109	7.3	430	1 SYH_CHLPN	Q927p1 chlamydia p
20	109	7.3	1520	1 GLTB_BACSU	P39812 bacillus su
21	109	7.3	1986	1 WA_EWEMI	Q03149 emericella
22	109	7.3	2334	1 WAPA_BACSU	Q07833 bacillus su
23	108.5	7.3	2477	1 FINC_MOUSE	P11276 mus musculu
24	108	7.3	1007	1 RGAL_YEAST	P39083 saccharomyc
25	107.5	7.2	329	1 PLC_BACTU	P08954 bacillus th
26	107.5	7.2	380	1 GLG2_YEAST	P47011 saccharomyc
27	107.5	7.2	429	1 DR48_YEAST	P18899 saccharomyc
28	107	7.2	420	1 YE09_YEAST	P40038 saccharomyc
29	107	7.2	1328	1 FINC_PLEVA	Q91289 pleurodeles
30	107	7.2	1562	1 YMB1_YEAST	Q04781 saccharomyc
31	106.5	7.2	493	1 GL03_YEAST	P38682 saccharomyc
32	106.5	7.2	843	1 YMS1_YEAST	Q05050 saccharomyc
33	106.5	7.2	2477	1 FINC_RAT	P04937 rattus norv

ALIGNMENTS

RESULT 1

ID	HLA_STAAU	STANDARD;	PRT;	319 AA.
AC	P09616;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Alpha-hemolysin precursor (Alpha-toxin) (Alpha-HL).			
GN	HLX OR HLA.			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=1280;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 27-44.			
RC	STRAIN=Wood 46;			
RX	MEDLINE=85053471; PubMed=6500704;			
RA	Gray G.S., Kehoe M.;			
RT	"Primary sequence of the alpha-toxin gene from Staphylococcus aureus wood 46.";			
RL	Infect. Immun. 46:615-618(1984).			
RN	[2]			
RP	REVISIONS, SEQUENCE FROM N.A.			
RC	STRAIN=Wood 46;			
RA	Kedungrahn G.;			
RN	Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.			
RC	SEQUENCE OF 27-319 FROM N.A., AND PARTIAL SEQUENCE.			
RX	STRAIN=Wood 46;			
RC	MEDLINE=92268149; PubMed=1587866;			
RA	Walker B., Krishnaswamy M., Zorn L., Kasianowicz J., Bayley H.;			
RT	"Functional expression of the alpha-hemolysin of Staphylococcus aureus in intact Escherichia coli and in cell lysates. Deletion of five C-terminal amino acids selectively impairs hemolytic activity.";			
RL	J. Biol. Chem. 267:10902-10909(1992).			
RN	[4]			
RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).			
RC	STRAIN=Wood 46 / ATCC 10832;			
RX	MEDLINE=97102581; PubMed=8943190;			
RA	Song L., Hobaugh M.R., Shustak C., Cheley S., Bayley H., Gouaux J.E.;			
RT	"Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore.";			
RL	Science 274:1859-1866(1996).			
RN	[5]			
RP	MUTAGENESIS.			
RC	MEDLINE=93016135; PubMed=1400487;			
RA	Walker B., Krishnaswamy M., Zorn L., Bayley H.;			
RT	"Assembly of the oligomeric membrane pore formed by Staphylococcal alpha-hemolysin examined by truncation mutagenesis.";			
RL	J. Biol. Chem. 267:21782-21786(1992).			
RN	[6]			
RP	MUTAGENESIS OF HISTIDINE RESIDUES.			
RC	MEDLINE=94222552; PubMed=8168947;			
RA	Menzies B.E., Kennodie D.S.;			
RT	"Site-directed mutagenesis of the alpha-toxin gene of Staphylococcus aureus: role of histidines in toxin activity in vitro and in a murine			

34	106	7.1	517	1	Z215_HUMAN	Q9ul58 homo sapien
35	106	7.1	537	1	TEB6_STRPY	P18481 streptococ
36	106	7.1	576	1	CSG_METVO	Q50833 methanococ
37	105.5	7.1	245	1	143B_RAT	P35213 rattus norv
38	105.5	7.1	928	1	HXA2_HAEIN	P45354 haemophilus
39	105	7.1	409	1	CAD1_YEAST	P24813 saccharomyc
40	104.5	7.0	638	1	GYRB_BACSU	P05652 bacillus su
41	104.5	7.0	655	1	CDGT_KLEPN	P08704 klebsiella
42	104.5	7.0	1145	1	POL_EIAVY	P03371 equine infe
43	104	7.0	311	1	PLC_STAAU	P45723 staphylococ
44	104	7.0	762	1	SLAP_ACEKI	P22258 acetogenium
45	104	7.0	928	1	NRPI_XENLA	P28824 xenopus lae

model.";
RL Infect. Immun. 62:1843-1847(1994).
[7]
RP MUTAGENESIS OF HISTIDINE RESIDUES.
RX MEDLINE=94245329; PubMed=8188346;
RA Jursch R., Hildebrand A., Hobom G., Tranum-Jensen J., Ward R.,
RA Kehoe M., Bhakdi S.;
RT "Histidine residues near the N-terminus of staphylococcal alpha-toxin
as reporters of regions that are critical for oligomerization and
pore formation.";
RL Infect. Immun. 62:2249-2256(1994).
[8]
RN MUTAGENESIS.
RP MEDLINE=96032742; PubMed=7559447;
RA Walker B., Bayley H.;
RT "Key residues for membrane binding, oligomerization, and pore forming
activity of staphylococcal alpha-hemolysin identified by cysteine
scanning mutagenesis and targeted chemical modification.";
RL J. Biol. Chem. 270:23065-23071(1995).
CC -!- FUNCTION: ALPHA-TOXIN BINDS TO THE MEMBRANE OF EUKARYOTIC CELLS
RESULTING IN THE RELEASE OF LOW-MOLECULAR WEIGHT MOLECULES AND
LEADING TO AN EVENTUAL OSMOTIC LYSIS. HEPTAMER OLIGOMERIZATION
AND PORE FORMATION IS REQUIRED FOR LYTIC ACTIVITY.
CC -!- SUBUNIT: SELF-ASSEMBLES TO FORM FIRST, A NONLYTIC OLIGOMERIC
INTERMEDIATE, AND THEN, A MUSHROOM-SHAPED HOMOPENTAMER STRUCTURE
OF 100 ANGSTROMS IN LENGTH AND UP TO 100 ANGSTROMS IN DIAMETER.
CC -!- SUBCELLULAR LOCATION: SECRETED AS A MONOMER. AFTER OLIGOMERIZATION
AND PORE FORMATION, THE COMPLEX IS TRANSLOCATED ACROSS THE
BILAYER, PROBABLY VIA THE GLY-RICH DOMAIN OF EACH STRAND.
CC -!- DOMAIN: THE MUSHROOM-SHAPED HEPTAMER IS COMPOSED OF A CAP DOMAIN
(COMPRISING 7 BETA SANDWICHES AND THE AMINO LATCHES OF EACH
PROTOMER), 7 RIM REGIONS WHOSE PROTRUDING STRANDS MAY INTERACT
WITH THE MEMBRANE BILAYER, AND THE STEM DOMAIN (52 ANGSTROMS IN
LENGTH, 26 ANGSTROMS IN DIAMETER) WHICH FORMS THE TRANSMEMBRANE
PORE.
CC -!- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: X01645; CAA25801.1; -;
DR EMBL: M90536; AAA26598.1; -;
DR PDB: 7AHL; 14-JAN-98.
DR InterPro: IPR001776; Aerolysin.
DR InterPro: IPR001340; Hemolysin_pore.
DR InterPro: IPR003963; Staph_bicn_txn.
DR Pfam: PF01117; Aerolysin; 1.
DR PRINTS: PR01468; BICOMPOTOXIN.
DR TIGRFAMs: TIGR01002; hlyII; 1.
DR PROSITE: PS00274; AEROLYSIN; 1.
KW Hemolysin; Toxin; Signal; 3D-structure.
FT SIGNAL 1 26
FT CHAIN 27 319 ALPHA-HEMOLYSIN.
FT DOMAIN 145 169 GLY-RICH.
FT SITE 159 160 CLEAVAGE OF MONOMERS BY PROTEINASE K
FT SITE 161 162 CLEAVAGE OF MONOMERS BY PROTEINASE K
FT SITE 164 165 CLEAVAGE OF MONOMERS BY PROTEINASE K
FT SITE 165 166 CLEAVAGE OF MONOMERS BY PROTEINASE K
FT SITE 161 61 H->L: NO OLIGOMERIZATION NOR HEMOLYTIC
FT MUTAGEN 61 61 ACTIVITY.
FT MUTAGEN 61 61 H->R: NO OLIGOMERIZATION NOR HEMOLYTIC
FT MUTAGEN 74 74 ACTIVITY.
FT MUTAGEN 170 170 H->L: 7% OF NORMAL HEMOLYTIC ACTIVITY.
FT MUTAGEN 170 170 H->L: 16% OF NORMAL HEMOLYTIC ACTIVITY.

FT MUTAGEN 285 285 H->L: 46% OF NORMAL HEMOLYTIC ACTIVITY.
SQ SEQUENCE 319 AA; 35904 MW; 6711C415DF7EBF30 CRC64;
Query Match 100.0%; Score 1484; DB 1; Length 319;
Best Local Similarity 93.2%; Pred. No. 1.3e-84;
Matches 273; Conservative 20; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ADSDNINIKTGTDTGSGNTVTGKTLVTDYDKENGHMKHVFYSFIDDKNHNKKLLVIRTKGT 60
Db 27 ADSDNINIKTGTDTGSGNTVTGKTLVTDYDKENGHMKHVFYSFIDDKNHNKKLLVIRTKGT 86
Qy 61 IAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEMXTXTXGX 120
Db 87 IAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEMXTXTXGX 146
Qy 121 NKNXTXDTXKXGLXGNXKXGTXKXVQDPFKTILSPDCKKVGKVIENNMYNQWG 180
Db 147 NGNVTGDDTGKIGGLIGANVSIGHTLKXVQDPFKTILSPDCKKVGKVIENNMYNQWG 206
Qy 181 PYDRDSNPNVGNQLFMKTRNGSKMAADNFDLPNKASSLLSGSFDFATVITMDRKASK 240
Db 207 PYDRDSNPNVGNQLFMKTRNGSKMAADNFDLPNKASSLLSGSFDFATVITMDRKASK 266
Qy 241 QQTNIIDVIYVRDDYQLHWTSTNNKGTNTKDKWTDSSERYKIDWEKEEMTN 293
Db 267 QQTNIIDVIYVRDDYQLHWTSTNNKGTNTKDKWTDSSERYKIDWEKEEMTN 319
RESULT 2
HLGB_STAAH STANDARD; PRT; 325 AA.
ID HLGB_STAAH AC 007226;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gamma-hemolysin component B precursor (H-gamma-1) (H-gamma-I).
GN HLGB OR SA2209 OR MW2344.
OS Staphylococcus aureus (strain N315),
OS Staphylococcus aureus (strain MW2), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879, 196620, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N315.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus.";
RL Lancet 357:1225-1240(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA.";
RL Lancet 359:1819-1827(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93138806; PubMed=8423103;
RA Cooney J.C., Kienle Z., Foster T.J., O'Toole P.W.;
RT "The gamma-hemolysin locus of Staphylococcus aureus comprises three
linked genes, two of which are identical to the genes for the F and S
components of leukocidin.";

```

RL Infect. Immun. 61:768-771(1993).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=99310396; PubMed=3075655;
RA Cooney J.C., Mulvey M., Arbuthnot J., Foster T.;
RT "Molecular cloning and genetic analysis of the determinant for gamma-
RL lysin, a two-component toxin of Staphylococcus aureus.";
RT J. Gen. Microbiol. 134:2179-2188(1988).
CC -1- FUNCTION: GAMMA-HEMOLYSIN CAUSES HEMOLYSIS IN RED BLOOD CELLS.
CC -1- SUBUNIT: CONSISTS OF TWO PROTEIN COMPONENTS (H-GAMMA-I AND -II).
CC H-GAMMA-I IS ALSO THE F SUBUNIT OF LEUKOCIDIN.
CC -1- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP003137; BAB43511.1; -.
CC EMBL; AP004830; BAB96209.1; -.
CC EMBL; L01055; AAA26639.1; -.
CC HSP; P09616; FAHL.
CC InterPro: IPR001776; Aerolysin.
CC InterPro: IPR001340; Hemlysn_pore.
CC InterPro: IPR003963; Staph_bicn_txn.
CC Pfam: PF01117; Aerolysin; 1.
CC PRINTS; PR01468; BICOMPPTOXIN.
CC TIGRFAMS; TIGR01002; hlyII; 1.
CC Hemolysis; Toxin; Signal; Complete proteome.
KW SIGNAL 1 25
FT CHAIN 26 325 GAMMA-HEMOLYSIN COMPONENT B.
FT SEQUENCE 325 AA; 36711 MW; 082999EB9339B6A0 CRC64;
SQ
Query Match 26.1%; Score 388; DB 1; Length 325;
Best Local Similarity 29.9%; Pred. No. 9.6e-17;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;
QY 9 TGTDTI-GSNTTV---KTGDLVITYDKENG-----MHKKVYVFSDIDKHNHKKLVIR 56
DB 21 SGTANAEGKITPVSVKVKDDKVTLYKTATADSDKFKISQILTFNFKDKSYDKDTLVK 80
QY 57 TGTIAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTXEXMXT- 115
DB 81 ATGNINSGFVKPNPDYDFSKLYGAKYNYSSQSNDSYVVDYAPKNQNEEFQVQNTL 140
QY 116 -XTGXNXXNTXDTXXKXGLXGNXSXGTXKXVQDFKFTILESPTDKKVGKVFNN 173
DB 141 GYTFGGDISISNGL---SGGLNG-NTAFSETINYKQESYRTLSRNTNYKNVGVGVEAHK 196
QY 174 MVNONGPYDRDSNNPVYGNQFMKTRNGSKAADNFDLPNKASSLLSSGSPDFATVIT 233
DB 197 IMNNGWPGYGRDSFHPYTYGNEFLAGQSSAYAGQNFIAQHOMPLLSRSNPNPEFLSVLS 256
QY 234 MDRKASKQOONTIDVIYERVRDDYQLHWTSTNWKGTNTKDKWTDRSSERYKIDWEKEEM 291
DB 257 -HRQDGAKKSKITVYQRENDLYQIRWNGFYWAGANYKNEKTRFKSTYIEDWENHKV 313
RESULT 3
LUKF_STAAU
ID LUKF_STAAU STANDARD; PRT; 323 AA.
AC P31715;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leukocidin F subunit precursor (Gamma-hemolysin, H-gamma-I subunit).
GN LUKF.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;

```

```

RN SEQUENCE FROM N.A., AND SEQUENCE OF 26-66.
RP STRAIN=MRSA NO. 4;
RX MEDLINE=92246945; PubMed=1575738;
RA Rahman A., Nariya H., Izaki K., Kato I., Kamio Y.;
RT "Molecular cloning and nucleotide sequence of leukocidin F-component
RL gene (lukF) from methicillin resistant Staphylococcus aureus.";
RL Biochem. Biophys. Res. Commun. 184:640-646(1992).
RN [2]
RP SEQUENCE OF 27-85 AND 322-323.
RC STRAIN=RIMD 310925;
RX MEDLINE=93223836; PubMed=8467905;
RA Kamio Y., Rahman A., Nariya H., Ozawa T., Izaki K.;
RT "The two Staphylococcal bi-component toxins, leukocidin and gamma-
RL hemolysin, share one component in common.";
RL FEBS Lett. 321:15-18(1993).
CC -1- FUNCTION: LEUKOCIDIN CAUSES CYTOTOXIC CHANGES IN POLYMORPHONUCLEAR
CC LEUKOCYTES. GAMMA-HEMOLYSIN CAUSES HEMOLYSIS IN RED BLOOD CELLS.
CC -1- SUBUNIT: LEUKOCIDIN CONSISTS OF TWO PROTEIN COMPONENTS: F AND S;
CC GAMMA-HEMOLYSIN CONSISTS OF TWO PROTEIN COMPONENTS (H-GAMMA-I-F
CC AND H-GAMMA-II).
CC -1- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S65052; AAC60446.1; -.
CC PIR; JQ1530; JQ1530.
CC InterPro: IPR001776; Aerolysin.
CC InterPro: IPR001340; Hemlysn_pore.
CC InterPro: IPR003963; Staph_bicn_txn.
CC Pfam: PF01117; Aerolysin; 1.
CC PRINTS; PR01468; BICOMPPTOXIN.
CC TIGRFAMS; TIGR01002; hlyII; 1.
CC Hemolysis; Toxin; Signal.
KW SIGNAL 1 25
FT CHAIN 26 323 LEUKOCIDIN F SUBUNIT.
FT SEQUENCE 323 AA; 36567 MW; A5951CF2AAB6C7E8 CRC64;
SQ
Query Match 25.2%; Score 374; DB 1; Length 323;
Best Local Similarity 29.3%; Pred. No. 7e-16;
Matches 87; Conservative 73; Mismatches 117; Indels 20; Gaps 8;
QY 9 TGTDTI-GSNTTV---KTGDLVITYDKENG-----MHKKVYVFSDIDKHNHKKLVIR 56
DB 21 SGTANAEGKITPVSVKVKDDKVTLYKTATADSDKFKISQILTFNFKDKSYDKDTLVK 80
QY 57 TGTIAGQYRVYSEEGANKSGLAWPSAFKVLQLPDNEVAQISDYYPNRSIDTXEXMXT- 115
DB 81 ATGNINSGFVKPNPDYDFSKLYGAKYNYSSQSNDSYVVDYAPKNQNEEFQVQNTL 140
QY 116 -XTGXNXXNTXDTXXKXGLXGNXSXGTXKXVQDFKFTILESPTDKKVGKVFNN 174
DB 141 GYTFGGDISISNGL---SGGLNG-NTAFSETINYKQESYRTLSRNTNYKNVGVGVEAHK 196
QY 175 VNQNWPGYDRDSNNPVYGNQFMKTRNGSKAADNFDLPNKASSLLSSGSPDFATVITM 234
DB 197 MN-GWPGYGRDSFHPYTYGNEFLAGQSSAYAGQNFIAQHOMPLLSRSNPNPEFLSVLS- 254
QY 235 DRKASKQOONTIDVIYERVRDDYQLHWTSTNWKGTNTKDKWTDRSSERYKIDWEKEEM 291
DB 255 HRQDRAKKSKITVYQRENDLYQIRWNGFYWAGANYKNEKTRFKSTYIEDWENHKV 311
RESULT 4
HLGA_STAAU
ID HLGA_STAAU STANDARD; PRT; 309 AA.
AC P31714; Q07225; Q53689; Q53690;

```



```
Db 200 SAYDOYLFAQDPTGP-AARDYFVPDNLQPLLIQSGFNPSFITTLSHERGKGDKE--FEI 256
QY 248 IYERVRDYQLHWTSTNWKCTNTKDKWTDERS-SERYKIDWEKEEM 291
Db 257 TYGRNMDATYAYTRHRLAVDRKDAFKNRNVTYKIEVNWKTHEV 301

RESULT 5
HLGC_STAAU
ID HLGC_STAAU STANDARD; PRT; 315 AA.
AC Q07227;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gamma-hemolysin component C precursor (H-gamma-1) (H-gamma-I).
GN HLGC.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93138806; PubMed=8423103;
RA Cooney J.C., Klenie Z., Foster T.J., O'Toole P.W.;
RT "The gamma-hemolysin locus of Staphylococcus aureus comprises three
RT linked genes, two of which are identical to the genes for the F and S
RT components of leukocidin."
RL Infect. Immun. 61:768-771(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89310396; PubMed=3075655;
RA Cooney J.C., Mulvey M., Arbutnot J., Foster T.;
RT "Molecular cloning and genetic analysis of the determinant for gamma-
RT lysin, a two-component toxin of Staphylococcus aureus."
RL J. Gen. Microbiol. 134:2179-2188(1988).
CC -1- FUNCTION: GAMMA-HEMOLYSIN CAUSES HEMOLYSIS IN RED BLOOD CELLS.
CC -1- SUBUNIT: CONSISTS OF TWO PROTEIN COMPONENTS (H-GAMMA-I AND -II).
CC H-GAMMA-I IS ALSO THE F SUBUNIT OF LEUKOCIDIN.
CC -1- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L01055; AAA26638.1; -.
CC InterPro; IPR001776; Aerolysin.
CC InterPro; IPR001340; Hemolysn_pore.
CC InterPro; IPR003963; Staph_bicn_tnx.
CC Pfam; PF01117; Aerolysin; 1.
CC PRINTS; PR01468; BICOMPFOXIN.
CC TIGRFAMS; TIGR01002; hlyII; 1.
CC Hemolysins; Toxin; Signal.
CC SIGNAL 1 29
CC CHAIN 30 315
CC SEQUENCE 315 AA; 35614 MW; 77359819736620BC CRC64;
CC -----
CC POTENTIAL.
CC SIGNAL 30 315
CC SEQUENCE 315 AA; 35614 MW; 77359819736620BC CRC64;
CC -----
Query Match 16.7%; Score 248; DB 1; Length 315;
Best Local Similarity 21.5%; Pred. No. 4.2e-08;
Matches 63; Conservative 85; Mismatches 117; Indels 28; Gaps 11;
QY 11 TTDIGSNTTVKTDGLV--TYDKEN---GMHKVFYFSDIDKHNKLLVTRTKGTAGQY 65
| ||| : : : | : : | : : | : : | : : | : : | : : | : : | : :
33 TEDIGKSGDI---EIIKRTEDKTSNKGVTQNIQFQVFKDKYNKDAILKMQGFISRT 89
| : : : : : | : : | : : | : : | : : | : : | : : | : : | : :
66 RVYSEEGANK-SGLWPSAFKVOQLPNDNEVAOISDYYPNSIDTXEMXTYTXGXNXX 124
| : : : : : | : : | : : | : : | : : | : : | : : | : : | : :
90 TYNYKNTNHVKARWPFQNIIGLTKNDKYSLI-NYLPKNKIESTN--VSQTLGNIG 146
| : : : : : | : : | : : | : : | : : | : : | : : | : : | : :
125 TXDXTKXGXGLXGXNXXSGXTKXKXQVPDFKLTILESPTDKKVGKVFNNMNQWGPYDR 184
```

```
Db 147 NFOSAPSLGNGSGFNYS--KSISYTOQNYVSEVEQNSKSVLWGVKANSFATES----- 198
QY 185 DSNVPYGNOLEFMKTRNGSKMKAADNFDLPNKASSLLSSGSPDFAITVITMDRKASKQOTN 244
Db 199 -GOKSAFDSLFFVGYKPHSKDPRDYFVPDSELPPLVOSGFNPFIATVSHE-KGSDTSE 256
QY 245 IDVIYERVRD-----DYQLHWTSTNWKGTNTKDKWTDERS-SERYKIDWEKEEM 291
Db 257 FEITYGRNMDVTTHAIKRSTHYGNSYLDGHRVHNAFVNRNVTYKIEVNWKTHEI 309

RESULT 6
LUKS_STAAU
ID LUKS_STAAU STANDARD; PRT; 315 AA.
AC P31716;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leukocidin S subunit precursor.
GN LUKS.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-79.
RX STRAIN=NO.4;
RX MEDLINE=92068180; PubMed=1958181;
RA Rahman A., Izaki K., Kato I., Kamio Y.;
RT "Nucleotide sequence of leukocidin S-component gene (lks) from
RT methicillin resistant Staphylococcus aureus."
RL Biochem. Biophys. Res. Commun. 181:138-144(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=MRSA NO. 4;
RX MEDLINE=93372503; PubMed=7763998;
RA Rahman A., Izaki K., Kamio Y.;
RT "Gamma-hemolysin genes in the same family with lukF and lks genes in
RT methicillin resistant Staphylococcus aureus."
RL Biosci. Biotechnol. Biochem. 57:1234-1236(1993).
CC -1- FUNCTION: LEUKOCIDIN CAUSES CYTOTOXIC CHANGES IN POLYMORPHONUCLEAR
CC LEUKOCYTES.
CC -1- SUBUNIT: LEUKOCIDIN CONSISTS OF TWO PROTEIN COMPONENTS: F AND S.
CC -1- SIMILARITY: TO THE H-GAMMA II SUBUNIT OF GAMMA-HEMOLYSIN.
CC -1- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M81346; AAA26654.1; -.
CC EMBL; S65052; AAC60445.1; -.
CC PIR; JN0626; JN0626.
CC InterPro; IPR001776; Aerolysin.
CC InterPro; IPR001340; Hemolysn_pore.
CC InterPro; IPR003963; Staph_bicn_tnx.
CC Pfam; PF01117; Aerolysin; 1.
CC PRINTS; PR01468; BICOMPFOXIN.
CC TIGRFAMS; TIGR01002; hlyII; 1.
CC Hemolysins; Toxin; Signal.
CC SIGNAL 1 29
CC CHAIN 30 315
CC SEQUENCE 315 AA; 35557 MW; 6F480166DF13AA1E CRC64;
CC -----
Query Match 16.3%; Score 242; DB 1; Length 315;
Best Local Similarity 21.2%; Pred. No. 1e-07;
Matches 62; Conservative 85; Mismatches 118; Indels 28; Gaps 11;
QY 11 TTDIGSNTTVKTDGLV--TYDKEN---GMHKVFYFSDIDKHNKLLVTRTKGTAGQY 65
```



```
Db 33 TEDIGKSDI---EIKRTEDTSKNGWVTONIQFVDRKTDALILKMQGFSSRT 89
QY 66 RYSEEGANK-SGLAWPSAFKVLQLPDNEVAQISDYPRNSIDIXEMXMTYXGNXKX 124
Db 90 TTYNYKTTNHRVAMRPFOYNTGLKTDNDYVSLI-NYLPKNKIESTN--VSOTLGYNIGG 146
QY 125 TXDXTKXGKXGNXSGXTKXKXVOPDFKTTILESPDKKVGKVIFFNNVNNQNWGPYDR 184
Db 147 NQOSAPSLGNGSFNYS--KSIYTOQNVVSEVQONSXVLWGVKANSFATES----- 198
QY 185 DSNPNVYGQLPMKTRNGSKMAADNPLDPNKASSLLSGFSPDFATVITMDRKASKQQTN 244
Db 199 -GQSAFSDSLFVGYPKSPDRYFVPDSELPPLVQSGFNPFSFIATVSHE-KGSDTSE 256
QY 245 IDVIVYERVD-----DYQLHWTSTNKKGTNTKDKWTDTRS-SERYKIDWEKEEM 291
Db 257 FEITYGRNMDVTHAIRKSTRTHYNSYLDGHRVHNAFVNRNRYTVKYEVNWKTHEI 309

RESULT 7
ID KYKL_DICDI STANDARD; PRT; 1584 AA.
AC P18160;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Non-receptor tyrosine kinase spore lysis A (EC 2.7.1.112) (Tyrosine-
protein kinase 1).
GN PYKA OR SPLA OR DPYK1.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Metazoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JH10;
RX MEDLINE=97053827; PubMed=8989241;
RA Nuckolls G.H., Oshrov N., Loomis W.F., Spudich J.A.;
RT "The Dictyostelium dual-specificity kinase spla is essential for
spore differentiation.";
RL Development 122:3295-3305(1996).
RN [2]
RP SEQUENCE OF 1248-1584 FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
Dictyostelium discoideum.";
RL Mol. Cell. Biol. 10:3578-3583(1990).
CC -!- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
tyrosine phosphate.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK
DURING THE MOUND STAGE OF MORPHOGENESIS.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 1 SAM DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; U32174; AAB41125.1; -
DR EMBL; M33785; AAA33202.1; -
DR PIR; A35670; A35670.
DR Dictydb; DD03010; pyka.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR003878; SPRY_domain.
DR InterPro; IPR003877; SPRY_receptor.
DR InterPro; IPR004040; STY_pkinase.
```

```
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR Pfam; PF00622; SPRY; 3.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00449; SPRY; 3.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00105; SAM_DOMAIN; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT DOMAIN 908 972 SAM.
FT DOMAIN 403 420 POLY-ASN.
FT DOMAIN 428 435 POLY-THR.
FT DOMAIN 449 480 POLY-ASN.
FT DOMAIN 483 491 POLY-ASN.
FT DOMAIN 494 508 POLY-ASN.
FT DOMAIN 512 532 POLY-ASN.
FT DOMAIN 596 600 POLY-ASN.
FT DOMAIN 808 811 POLY-PHE.
FT DOMAIN 1026 1029 POLY-SER.
FT DOMAIN 1195 1210 POLY-ASN.
FT DOMAIN 1215 1220 POLY-GLN.
FT DOMAIN 1224 1233 POLY-GLN.
FT DOMAIN 1266 1274 POLY-PRO.
FT DOMAIN 1289 1561 PROTEIN_KINASE.
FT NP_BIND 1295 1303 ATP (BY SIMILARITY).
FT BINDING 1316 1316 ATP (BY SIMILARITY).
FT ACT_SITE 1417 1417 D -> R (IN REF. 2).
FT CONFLICT 1248 1248 V -> L (IN REF. 2).
FT CONFLICT 1435 1435
SQ SEQUENCE 1584 AA; 174304 MW; 5D1589458D8E01E3 CRC64;

Query Match 8.0%; Score 118; DB 1; Length 1584;
Best Local Similarity 20.7%; Pred. No. 60;
Matches 36; Conservative 48; Mismatches 50; Indels 40; Gaps 6;

QY 93 NEVAQI-----SDYPRNSIDTXEMXMTYXGNXNXTXDTKXKXGKXNXXGXTXK 147
Db 967 NRMIOIWNKSPDYPKTAIDSSDKIRWPASGGSGG---INISGVVIGSSGDDGIT 1023
QY 148 XVQPDFKTI--LESPTDKKVGKVIFFNNVNNQNWGPYDRDSNPNVYG----NOLFMTNRN 201
Db 1024 EISSSSKNIRPVKSYTQKEI-----EDNRNSTISGGEKKNKYI---- 1063
QY 202 GSKAAADNPLDPNKASSLLSGFSPDFATVITMDRKASKQQTNIDVIVYERVD 255
Db 1064 -----DNQMDPHQIGSMDSGLLPDFGQPPDEKNSKTLTNEQIRYLQQRKD 1111

RESULT 8
CHD3_CAEBL STANDARD; PRT; 1787 AA.
ID CHD3_CAEBL
AC Q22516; Q18794;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chromodomain helicase-DNA-binding protein 3 homolog (CHD-3).
GN CHD-3 OR TL468.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20530482; PubMed=11076750;
RA von Zelewsky T., Palladino F., Brunschwig K., Tobler H., Hajnal A.,
RA Mueller F.;
RT "The C. elegans Mi-2 chromatin-remodelling proteins function in vulval
cell fate determination.";
RL Development 127:5277-5284(2000).
```



```
FT CHAIN 76 546 MJA RPOA2 INTEIN (SPACER PROTEIN)
FT CHAIN 547 859 RPOA2, 2ND PART (POTENTIAL)
FT CHAIN 547 859 RPOA2, 2ND PART (POTENTIAL)
SQ SEQUENCE 859 AA; 97086 MW; A03799E007899015 CRC64;

Query Match 7.8%; Score 116; DB 1; Length 859;
Best Local Similarity 23.6%; Pred. No. 31;
Matches 58; Conservative 53; Mismatches 99; Indels 36; Gaps 13;

QY 26 VYDKENGHHKVFYFIDDKHKKLLVIRKGV---TIAGQYRVYSEGGANK-----S 76
DB 127 LSLDQDEKQVHWRKLIISCIHKK-HNGKLIKIKTKSGREITATPTPHSVIRKDKNIIPVKGS 185

QY 77 GLAWPSAFVQLQLPDNP--EVAQISDYVPNSI--DTXEXMXTXTXGXNKNXTXDXTKX 132
DB 186 ELKIGDRIPVKKHIPANCVAINISYVSGNIVVDNINNKIAPKNGKIPNNIKLIDYDF 245

QY 133 GXLGXNXXSGXTXKXVQDPFKTILESPTDKKVGKVV-IFNNMNVNONGPYPDSDSNWP-- 189
DB 246 GYFIGIYLAEGSVTKY---FVSI--SNVDEILINKIRAFADKGLGLNYGEYDNNNGFAES 299

QY 190 ---VYGNOL--FMK---TRNSGMAKADNFDLPNK--ASSLLSGSPSPFAVITMDRKA 238
DB 300 HDIRYSSSTLAELFSLNFGTSSNTTKIAEFVGCANKEFVRGLIRGYPDGD--GNVNA DRKV 357

QY 239 SKOOTN 244
DB 358 IRVTSN 363

RESULT 11
EPIP_STAMP STANDARD; PRT; 461 AA.
AC P30199.
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epidermin leader peptide processing serine protease EPIP precursor
DE (EC 3.4.21.-).
GN EPIP.
OS Staphylococcus epidermidis.
OG Plasmid pTu 32.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TU 3298 / DSM 3095;
RX MEDLINE=92155237; PubMed=1740156;
RA Schnell N., Engelke G., Augustin J., Rosenstein R., Ungermann V.,
RA Goetz F., Entian K.-D.;
RT "Analysis of genes involved in the biosynthesis of lantibiotic
RT epidermin."
RL Eur. J. Biochem. 204:57-68(1992).
CC -!- FUNCTION: PROTEASE WHICH CLEAVES THE MATURED LANTIBIOTIC FROM THE
CC MODIFIED PREPEPTIDE (PROBABLE).
CC -!- PATHWAY: Epidermin biosynthesis; last step.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X62386; CAA44257.1; -.
CC PIR; S23420; S23420.
CC HSP; P00782; 2SBT.
CC MEROPS; S08.060; -.
CC InterPro; IPR000209; Peptidase_S8.
CC Pfam; PF00082; Peptidase_S8; 1.
CC PRINTS; PR00723; SUBTILISIN.
```

```
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
KW Hydrolase; Serine protease; Signal; Zymogen; Plasmid.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 ?
FT CHAIN ? 461
FT EPIDERMIN LEADER PEPTIDE PROCESSING
FT SERINE PROTEASE EPIP.
FT ACT_SITE 149 149 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 194 194 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 402 402 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 461 AA; 51814 MW; F2770F4F8436D906 CRC64;

Query Match 7.7%; Score 115; DB 1; Length 461;
Best Local Similarity 21.7%; Pred. No. 13;
Matches 63; Conservative 57; Mismatches 118; Indels 52; Gaps 13;

QY 4 DINIKGTTDIGS-----NTTVKTDGLVTYDKENGHMKKVFYFIDDKHNNKLLV--IR 56
DB 149 DSGVNSSHTDLKSINKIVNEVPKNGFRGSEDNESGK-----NFEEDKLNHGTLVAGQIG 203

QY 57 TKGTTAG-----QYRVYSEGGANKSGLAWPSAFKVLQLPDNEVAQIS-----DYYP 104
DB 204 ANGNLKGVNPGVEMNMYRVFVSGSKSEMLWVSKGIIDAANDNDVINVSLGNVLIKDNQNK 263

QY 105 NSIDTXEXMXTXTGXNKNXTXKXGLXGXNXXSGXTXKXVQ-PDFKFILESPTDK 163
DB 264 KKLDRDEKVDYDALOKAINTAQ---KKSIVVAAVGNCGDINVKKKEINKRNLSKTSK 320

QY 164 KVGWKVIFNNMVNQN---WGYPYDRDSNWPV---YGNQLF-MKTRNGSMKAADNFDLPNK 215
DB 321 KV-----YDSPANLNVMVTGSDIDNDYISEFSNYGNFNIDLMTIGSYK----LLDKYG 371

QY 216 ASSLLSGSPSPFAVITMDRASKQQTWIDVIYRVR-----DDYQL 258
DB 372 KQAWLEKGYMQKQSVLSTSSNGRIYQSGTSLAAPKVGALALEIDKYQL 421
DE (Royal jelly protein RJP57-2).
GN MRJP4.
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC Aculeata; Apoidea; Apidae; Apis.
OX NCBI_TaxID=7460;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RA Klaudiny J., Hanes J., Kullafajova J., Albert S., Slmuth J.;
RT "Molecular cloning of two cDNAs from the head of the nurse honey bee
RT (Apis mellifera L.) for coding related proteins of royal jelly.";
RL J. Apicultural Res. 33:105-111(1994).
RN [2]
RC CHARACTERIZATION.
RC TISSUE=Head;
RX MEDLINE=99007754; PubMed=9791542;
RA Hanes J., Judova J., Slmuth J.;
RA "A family of major royal jelly proteins of the honeybee Apis mellifera
RA L.";
RL Cell. Mol. Life Sci. 54:1020-1030(1998).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN HONEYBEE NUTRITION. IT IS
CC FOUND IN THE ROYAL JELLY WHICH IS THE FOOD OF THE QUEEN HONEY BEE
CC LARVA. THE ROYAL JELLY DETERMINES THE DEVELOPMENT OF THE YOUNG
CC LARVAE AND IS RESPONSIBLE FOR THE HIGH REPRODUCTIVE ABILITY OF THE
```



```

DR EMBL; AL008725; CAA15497.1; -
DR EMBL; BC001359; AAH01359.1; -
DR HSSP; P29312; 1A38.
DR Genew; HGNC:12849; YWHAB.
DR MIM; 601289; -
DR InterPro; IPR000308; 14-3-3.
DR Pfam; PF00244; 14-3-3; 1.
DR PRINTS; PR00305; 1433ZETA.
DR ProDom; PD000600; 14-3-3; 1.
DR SMART; SM00101; 14-3-3; 1.
DR PROSITE; PS00796; 1433_1; 1.
DR PROSITE; PS00797; 1433_2; 1.
KW Brain; Neurone; Phosphorylation; Acetylation; Multigene family;
KW Alternative initiation.
FT INIT_MET 0
FT CHAIN 1 245
FT CHAIN 2 245
FT INIT_MET 2 2
FT MOD_RES 1 1
FT MOD_RES 2 2
FT MOD_RES 185
FT SEQUENCE 245 AA; 27951 MW; 0BCA59BF97595485 CRC64;
Query Match 7.6%; Score 113.5; DB 1; Length 245;
Best Local Similarity 19.1%; Pred. No. 6.1;
Matches 50; Conservative 52; Mismatches 87; Indels 73; Gaps 9;
QY 28 YKENGKMKVYFFDIDKHNKLLVIRTKGTIAQ---YRVYS-----BEGANKSL 78
DB 20 YDMAAMKAVTEQGHLSNEERLLSVAYKNVVGARRSWRVSSIEOKTERNEKKQOM 79
QY 79 AMPFAKVLQLPD--NEVAQISDYI-----PRNSI-----DTXEMXTXTGXNX 122
DB 80 GREYREKIEALQDNDVLELLDKYLIPNATQPSKVYLMKMGDYFYLSEVASGDNK 139
QY 123 NYTXDYTKXGXLGXNKSAGTXXVQDPFTILESPDQKVGKVFIPNNMVNONWGPY 182
DB 140 QTTVSNQQA-----YQEAFFISKEMQPTPIRLGLALNFSVFYIELNS----- 185
QY 183 DRDSWNPVYGNLFMKTRNGSKAADNFDPNKASLLSSGFSPOFATVITMDRKASKQ 242
DB 186 -----PEKACSLAKTAFDEATAELDTLNEESYKDS 215
QY 243 TNIDVIYERVDYQLHWTSTN 264
DB 216 T---LIMQLLRNLTL-WTSEN 233
RESULT 14
CNA_STAAU STANDARD; PRT; 1183 AA.
AC Q53554;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Collagen adhesin precursor.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FDA 574;
RX MEDLINE=92165839; PubMed=1311320;
RA Patti J.M., Jonsson H., Guss B., Switalski L.M., Wiberg K.,
RA Lindberg M., Hoeek M.;
RT "Molecular characterization and expression of a gene encoding a
RT Staphylococcus aureus collagen adhesin.";
RL J. Biol. Chem. 267:4766-4772(1992).
RN [2]
RP ERRATUM.
RA Patti J.M., Jonsson H., Guss B., Switalski L.M., Wiberg K.,

```

```

RA Lindberg M., Hoeek M.;
RL J. Biol. Chem. 267:11672-11672(1994).
RN [3]
RP COLLAGEN-BINDING DOMAIN.
RC STRAIN=FDA 574;
RX MEDLINE=94032261; PubMed=8218209;
RA Patti J.M., Boles J.O., Hoeek M.;
RT "Identification and biochemical characterization of the ligand
RT binding domain of the collagen adhesin from Staphylococcus aureus.";
RL Biochemistry 32:11428-11435(1993).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 169-318.
RX MEDLINE=97475225; PubMed=9334749;
RA Symersky J., Patti J.M., Carson M., House-Pompeo K., Teale M.,
RA Moore D., Jin L., Schneider A., Delucas L.J., Hoeek M.,
RA Narayana S.V.L.;
RT "Structure of the collagen-binding domain from a Staphylococcus
RT aureus adhesin.";
RL Nat. Struct. Biol. 4:833-838(1997).
CC -!- FUNCTION: MEDIATES ATTACHMENT OF STAPHYLOCOCCAL CELLS TO
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; M81736; AAA20874.1; -
DB PDB; 1AMX; 24-JUN-98.
DR InterPro; IPR001899; Gram_pos_anchor.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; FALSE_NEG.
KW Cell wall; Peptidoglycan-anchor; Repeat; Signal; 3d-structure.
FT SIGNAL 1 29
FT CHAIN 30 1154
FT PROPEP 1155 1183
FT DOMAIN 151 318
FT DOMAIN 533 1093
FT DOMAIN 1093 1157
FT REPEAT 533 719
FT REPEAT 720 906
FT REPEAT 907 1093
FT SITE 1151 1155
FT MOD_RES 1154 1154
FT SEQUENCE 1183 AA; 133066 MW; B6A1CC072E575D76 CRC64;
Query Match 7.5%; Score 112; DB 1; Length 1183;
Best Local Similarity 18.9%; Pred. No. 90;
Matches 60; Conservative 62; Mismatches 118; Indels 78; Gaps 12;
QY 3 SDINIKTGTDTIGSNTVTGTGLVTVYDKENGHMKKVFYFIDDKHNKLLVIRTKGTIA 62
DB 39 TDLTVSPSKIEDGGKTVK---MTFDDKNG-----KIQGDMIKV----- 75
QY 63 GOYRVYSEEGANKSGLAWPSAFKVLQI-----LPDNEVAQISDYI 102
DB 76 -----AWPTSGTVKIEGYSKTVPLTVKGEQGVQAVITPDGATITFNDKV 119
QY 103 PRNSIDTXXEMXTXTGXNKNXTXDTXKXGXLGXNKSAGTXXVQDP-----FKTI 156
DB 120 EKLK-DVSGFAEFVQGRNLQTNTSDDKVATITSGNKSTNTVHKSEAGTSVFYKYG 178
QY 157 LESPTD-KVGVKVIFFNNMVNONWGPYDRDSWNPVY-GNOLFMTKTRNGSKAA-DNFLDP 213
DB 179 DMLPETHVRFNLINN--EASVYKSDITIKDQIQGGQQLDLSTLINIVTTHSNYSG 236
QY 214 NKASLLSSGFSPOFATVITMDRKASKQOOTNIDVIYERVDYQLHWTSTNKGNTKDK 273

```

```
Db 237 QSAITDFEKAFF---GSKITVD-----NFKNTIDVTIPOGYSYNSFSINYTKITNEQOK 289
Qy 274 WTDSSERYKIDWEKEEM 291
Db 290 EFVNSQAWQEHGKEEV 307

RESULT 15
TFC3_YEAST
ID TFC3_YEAST STANDARD; PRT; 1160 AA.
AC P34111;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE Transcription factor tau 138 kDa subunit (TFC3 138 kDa subunit).
GN TFC3 OR TSV115 OR YAL001C OR FUN24.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=S288C;
RX MEDLINE=93066269; PubMed=1279682;
RA Lefebvre O., Carles C., Conesa C., Swanson R.N., Bouet F., Riva M.,
RA Sentenac A.;
RT "TFC3: gene encoding the B-block binding subunit of the yeast
RT transcription factor IIIC.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10512-10516(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=95028152; PubMed=7941740;
RA Clark M.W., Keng T., Storms R.K., Zhong W., Fortin N., Zeng B.,
RA Delaney S., Ouellette B.F.F., Barton A.B., Kaback D.B., Bussey H.;
RT "Sequencing of chromosome I of Saccharomyces cerevisiae: analysis of
RT the 42 kbp SPO7-CEN1-CDC15 region.";
RL Yeast 10:535-541(1994).
CC -1- FUNCTION: TFC3 MEDIATES TRNA AND 5S RNA GENE ACTIVATION BY
CC BINDING TO INTRAGENIC PROMOTER ELEMENTS. IT ASSEMBLES THE
CC INITIATION COMPLEX TFC3-TRNA. TFC3 IS ESSENTIAL FOR
CC CELL VIABILITY.
CC -1- SUBUNIT: COMPOSED OF AT LEAST FOUR DIFFERENT SUBUNITS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: SOME, TO S.POMBE SPBC336.07.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M98261; AAA34378.1; -.
CC EMBL; L22015; AAC04956.1; -.
CC PIR; A46423; A46423.
CC PIR; S40892; S40892.
CC PIR; S43444; S43444.
CC TRANSFAC; T03679; -.
CC SGD; S0000001; TFC3.
CC Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 1 68 HMG BOX.
FT DNA_BIND 1037 1110 HMG BOX.
FT SEQUENCE 1160 AA; 132108 MW; 3ACB7893ED536581 CRC64;

Query Match 7.5%; Score 111.5; DB 1; Length 1160;
Best Local Similarity 16.1%; Pred. No. 93;
Matches 66; Conservative 83; Mismatches 114; Indels 147; Gaps 15;

Qy 1 ADSDINKTGTDTGSGNTVTGTLVTDKEN--GMHKVFYFIDD----- 45
Db 418 AEDEISVPKGFDELGKSR-----DLKTLNEDNFVALNTRVFTTDSGDQDIFFWHGELKI 473
```

```
Qy 46 -----KNHKKLLVIRTKGIAGQYRVYSEEGANKSLWAPSAFKVOLQLPDNEVA 96
Db 474 PPNSKKTTPNKNRRQVKNSTNASVAG-----NISPKRIKLEQHVSTAQEP 520
Qy 97 QISDYYPNRSIDT-XEXMTXTXGXNXXNTXDTXKXGLXGXNXXSGXT----- 145
Db 521 KSAEDSPSSNGGTVVKGKVVNFGFSARSLRSLQRRAILKVMNTIGGVAYLREQFYEV 580
Qy 146 -----XKXVQDPFKTILESP-----TDKKVGVKVIENNMVN-----QNHGPDYDRD 185
Db 581 SKYMGSTTTTLDKKTVRGVDVLDLWVESEKLGARTPEVSGRKIIIFLPTVGEDAIQRIKED 640
Qy 186 S-----WNPVYGNOLEM-----KTRNGSMKAADNFDPNKA 216
Db 641 SKKATFTDVIHDTIYFEDQTEKRNFRHGKKSVERIRKFNQRQNAKIKASDDAISKKST 700
Qy 217 SSLSSGFSPPDFATVITMDRKASKQOQNTIDVIYERVRDDYOLHWTST----- 263
Db 701 SVNVDG-----KIKRRDKKVSAGRTT--VVVENTKEDKTVYHAGTKDGVQALIRAVV 752
Qy 264 -----NWKGTNTKDKWTDR-----SSERYKID-WEK 288
Db 753 TKSINKIMWDKITKLPNNSLDNLKXKKTARRVYRMGHSGHWRAYVDKWK 802
```

Search completed: March 4, 2003, 10:43:08
Job time : 20 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2003, 10:40:04 ; Search time 28.5 seconds
(without alignments)
2118.309 Million cell updates/sec

Title: CLAIM7

Perfect score: 1484

Sequence: 1 ADSIDINIKGTDTIGSNVTY.....WDRSSRYKIDWEKEETN 293

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1480	99.7	319	16 Q99UU6	Q99uu6 staphylococ
2	437.5	29.5	412	2 O05387	O05387 bacillus ce
3	401	27.0	325	2 O50604	O50604 staphylococ
4	397	26.8	325	9 O80067	O80067 staphylococ
5	396	26.7	327	2 O54082	O54082 staphylococ
6	395.5	26.7	327	2 Q93UU8	Q93uu8 staphylococ
7	394.5	26.6	327	16 Q99T54	Q99t54 staphylococ
8	394.5	26.6	332	16 Q931N4	Q931n4 staphylococ
9	394	26.5	325	2 Q53747	Q53747 staphylococ
10	393	26.5	325	2 Q57227	Q57227 staphylococ
11	385	25.9	325	16 Q931F3	Q931f3 staphylococ
12	385	25.9	326	16 Q99SN8	Q99sn8 staphylococ
13	384.5	25.9	326	2 Q54327	Q54327 staphylococ
14	374	25.2	322	2 Q53704	Q53704 staphylococ
15	374	25.2	322	9 Q9MBN2	Q9mbn2 staphylococ
16	351	23.7	336	2 Q9EVAL	Q9eval bacillus ce

17	314	21.2	265	2 Q937V1	Q937v1 bacillus ce
18	313	21.1	265	2 Q937V0	Q937v0 bacillus ce
19	308.5	20.8	260	2 Q937V2	Q937v2 bacillus ce
20	293	19.7	336	2 Q46308	Q46308 clostridium
21	293	19.7	336	2 Q46181	Q46181 clostridium
22	289	19.5	309	2 Q9L403	Q9l403 clostridium
23	265	17.9	308	2 Q53731	Q53731 staphylococ
24	262	17.7	308	2 Q53703	Q53703 staphylococ
25	262	17.7	308	9 Q9MBN3	Q9mbn3 staphylococ
26	256.5	17.3	351	16 Q931I5	Q931i5 staphylococ
27	254.5	17.1	351	16 Q99SN7	Q99sn7 staphylococ
28	251	16.9	315	2 Q53701	Q53701 staphylococ
29	249	16.8	286	2 Q9AFB9	Q9afb9 staphylococ
30	248	16.7	286	2 Q9AFB0	Q9afb0 staphylococ
31	248	16.7	315	2 Q53691	Q53691 staphylococ
32	248	16.7	315	16 Q99RL1	Q99rl1 staphylococ
33	247.5	16.7	311	2 Q93UU9	Q93uu9 staphylococ
34	247.5	16.7	311	16 Q99T53	Q99t53 staphylococ
35	239	16.1	314	2 O54081	O54081 staphylococ
36	237	16.0	312	2 O50603	O50603 staphylococ
37	237	16.0	312	9 O80066	O80066 staphylococ
38	231	15.6	312	2 O53746	O53746 staphylococ
39	203.5	13.7	310	2 Q54326	Q54326 staphylococ
40	147	9.9	28	2 P97144	P97144 staphylococ
41	129	8.7	1272	16 Q98PR8	Q98pr8 mycoplasma
42	127.5	8.6	1348	5 O8T234	O8t234 dictyosteli
43	125.5	8.5	290	5 O94482	O94482 dictyosteli
44	124	8.4	410	9 Q92XE1	Q92xel bacterioph
45	119	8.0	447	2 Q9RPL5	Q9rpl5 streptococ

ALIGNMENTS

RESULT 1

Q99UU6 PRELIMINARY; PRT; 319 AA.
AC Q99UU6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Alpha-hemolysin precursor.
GN SAV1163 OR SA1007.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
Sekizizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus";
RL Lancet 357:1225-1240(2001).
DR ENBL; AP003361; BAB57325.1; -;
DR EMBL; AF003132; BAB42258.1; -;
DR HSSP; P09616; 7AHL.
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin, I.
DR PRINTS; PR01468; BICOMPFOXIN.
DR TIGRFAMS; TIGR01002; hlyII; 1.
DR PROSITE; PS00274; AEROLYSIN; 1.
KW Complete proteome.

```
SQ SEQUENCE 319 AA; 35975 MW; 9AB4C82FB2CD9698 CRC64;
Query Match 99.7%; Score 1480; DB 16; Length 319;
Best Local Similarity 92.8%; Pred. No. 3.8e-80;
Matches 272; Conservative 21; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADSDNINIKTGTDTIGSNWTTVKTGDLVTDKENGHMKVYFSFIDDKHNNKLLVIRTKGT 60
Db 27 ADSDNINIKTGTDTIGSNWTTVKTGDLVTDKENGHMKVYFSFIDDKHNNKLLVIRTKGT 86
QY 61 IAGQRYVSEEGANKSLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXXMXXTXGX 120
Db 87 IAGQRYVSEEGANKSLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXXMXXTXGX 146
QY 121 NXNXTYDXTXKXGXLGXNXXSXTKXXYQDPDKTILESPTDKKVGKVIFFNNMWNQNW 180
Db 147 NGNVTGDDTKIGGLIGANYSIGHTLKYQDPDKTILESPTDKKVGKVIFFNNMWNQNW 206
QY 181 PYDRDSNPNVYGNQLEFMKTRNGSKAADNFDLPNKASSLLSGFSDFATVITMDRKASK 240
Db 207 PYDRDSNPNVYGNQLEFMKTRNGSKAADNFDLPNKASSLLSGFSDFATVITMDRKASK 266
QY 241 QQTNIIDIVYVRDDYQLHWTSTNWKGTNTKDKWTDSSERYKIDWEKEEMTN 293
Db 267 QQTNIIDIVYVRDDYQLHWTSTNWKGTNTKDKWTDSSERYKIDWEKEEMTN 319

RESULT 2
O05387 PRELIMINARY; PRT; 412 AA.
AC O05387;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hemolysin II.
GN HLY-II.
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VKM-B771;
RX MEDLINE=94111714; PubMed=8283976;
RA Sinev M.A., Budarina Z.I., Gavrilenko I.V., Tomashevskii A.I.,
RA Kuzmin N.P.;
RT "Evidence of the existence of hemolysin II from Bacillus cereus:
cloning the genetic determinant of hemolysin II."
RL Mol. Biol. (Mosk.) 27:1218-1229(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-VKM-B771.
RX MEDLINE=20016563;
RA Baida G., Budarina Z.I., Kuzmin N.P., Solonin A.S.;
RT "Complete nucleotide sequence and molecular characterization of
hemolysin II gene from Bacillus cereus."
RL FEMS Microbiol. Lett. 180:7-14(1999).
DR EMBL; U94743; AAB51536.1; -
DR HSSP; P09616; 7AHL.
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemolysin_pore.
DR Pfam; PF01117; Aerolysin; 1.
DR TIGRFAMs; TIGR01002; hlyII; 1.
SQ SEQUENCE 412 AA; 45664 MW; D7E8313151C85A8F CRC64;

Query Match 29.5%; Score 437.5; DB 2; Length 412;
Best Local Similarity 31.9%; Pred. No. 7.6e-18;
Matches 88; Conservative 71; Mismatches 96; Indels 21; Gaps 7;

QY 27 TYDENGHMKVYFSFIDDKHNNKLLVIRTKGT-IAGQRYVSEEGANKSLAWPSAFK 85
Db 53 TYDMKNIKNSIKVSFIEDPYADKKIAIVTDDGSNIDAKYTI--NGGYNAGLWPSAYH 110
```

```
QY 86 VQLQLPDNEVAQISDYPRNSIDTXXMXXTXGXNXXTDXTKXKXGL-XGXNXSGX 144
Db 111 TEAEITSGDSAQFHKAAPVN-----TMTSAKYTSEVGYTLGSKVGVNDKGFNADASI 164
QY 145 TXXXV-----OPDFKTTLESPTDKKVGKVIFFNNMWNQNWGPYDRDSNPNVYGNQLF 196
Db 165 TGSFAWKESYSDQVDYKTVLETHTDKLNKVGFSQSFNPEWGIYNRDSFNTFYGNOLF 224
QY 197 MKYRNGSKMAADNFDLPNKASSLLSGFSDFATVITMDRKASKQQTNIIDIVYVRDDY 256
Db 225 MKRSYN-EGTNFVSKDTPVALTGFSNVAIVATDK--TESTSLKATNRRISDQY 281
QY 257 QLHWTSTNWKGTNTKDKWTDSSERYKIDWEKEEMT 292
Db 282 NIEWSSKMGWGTNNKDTYNEFFTNKYKLDWKNHQVT 317

RESULT 3
O50604 PRELIMINARY; PRT; 325 AA.
AC O50604;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE LukF-PV.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC49775;
RX MEDLINE=98067870; PubMed=9404084;
RA Kaneko J., Muramoto K., Kamio Y.;
RT "Gene of LukF-PV-like component of Panton-Valentine leukocidin in
Staphylococcus aureus P83 is linked with lukM."
RL Biosci. Biotechnol. Biochem. 61:1960-1962(1997).
DR EMBL; AB006796; BAA24008.1; -
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemolysin_pore.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPNTOXIN.
DR TIGRFAMs; TIGR01002; hlyII; 1.
SQ SEQUENCE 325 AA; 36958 MW; 195222EA2DA498424 CRC64;

Query Match 27.0%; Score 401; DB 2; Length 325;
Best Local Similarity 28.9%; Pred. No. 7.5e-16;
Matches 81; Conservative 72; Mismatches 107; Indels 20; Gaps 6;

QY 11 TTDIGSNWTTVKTGDLVTDKENGHMKVYFSFIDDKHNNKLLVIRTKGTIAGQRYVSE 70
Db 45 TTATSDSKLISKILIT-----FNFIKDSYDKDTLILKAAGNIVSGTKPNP 92
QY 71 EGANKSLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXXMXXTXGXNXXTDXTKX 130
Db 93 KDTISSQFYWGSKYNISINSDSNVNVVDYAPKN--QNEEFQVQQTGYSGV--GDINI 148
QY 131 KXGXLGXN--XSXGTXKXXYQDPDKTILESPTD-KYGVKVIFFNNMWNQNWGPYDRDSW 187
Db 149 SNGLSGGGNGSKSFSETINRYKQESYRTSLDKRTNFKKIGWDVEAHKIMNNGWPGYGRDSY 208
QY 188 NPVYGNQLFMKTRNGSKAADNFDLPNKASSLLSGFSDFATVITMDRKASKQQTNIIDV 247
Db 209 HSTGNEFMFLGSRNLAGQNFLEYHKMPVLVSRGNFPEFGLVLSRKNQNAK-KSKITV 267
QY 248 IYERVDDYQLHWTSTNWKGTNTKDKWTDSSERYKIDWE 287
Db 268 TYQREMDRYTNFWINENIGNNKNYKDNENPATHTSIYEDWE 307

RESULT 4
```



```
FT SIGNAL 1 26 POTENTIAL.
SQ SEQUENCE 327 AA; 36889 MW; 3807DB6421ACDAAA CRC64;

Query Match      26.7%; Score 395.5; DB 2; Length 327;
Best Local Similarity 30.0%; Pred. No. 1.9e-15;
Matches 82; Conservative 70; Mismatches 112; Indels 9; Gaps 6;

QY 18 TTVKTGDLVTDKENGMMHKVFYSFDDKNHKKLLVIRTKGTIAGQYRVYSEEGANKSG 77
Db 43 TLYKTTATSDNDKLN-ISOILTFNFIKDKSYDKDTLVLAAGNINSYGKKPNKDYNSQ 101

QY 78 LAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEXMXTYTXGXNXXTXDXTKXGXKXG 137
Db 102 FYWGGKYNVSVSESNDVAVVYAPKN--QNEEFQVQOTLGSYG--GDINISNGLSGG 157

QY 138 XN--XSXGXTXKXVQDPDKTILESPTD-KKVGWKVIFNNMNVNQNMGCPYDRDSWNPVYQ 194
Db 158 LSGKSFSTINYNKQESYRTTIDRTNHNKSGVGEAHKIMNMGWGPYGRDSYDPTYGNE 217

QY 255 DYQLHWTSTNWKGTNTKDKWDRSSERYKIDWE 287
Db 277 RYTQWNLRLHWGNKYNKQNTVTFSTVEVDWQ 309

QY 138 XN--XSXGXTXKXVQDPDKTILESPTD-KKVGWKVIFNNMNVNQNMGCPYDRDSWNPVYQ 194
Db 158 LSGKSFSTINYNKQESYRTTIDRTNHNKSGVGEAHKIMNMGWGPYGRDSYDPTYGNE 217

QY 18 TTVKTGDLVTDKENGMMHKVFYSFDDKNHKKLLVIRTKGTIAGQYRVYSEEGANKSG 77
Db 43 TLYKTTATSDNDKLN-ISOILTFNFIKDKSYDKDTLVLAAGNINSYGKKPNKDYNSQ 101

QY 78 LAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEXMXTYTXGXNXXTXDXTKXGXKXG 137
Db 102 FYWGGKYNVSVSESNDVAVVYAPKN--QNEEFQVQOTLGSYG--GDINISNGLSGG 157

QY 138 XN--XSXGXTXKXVQDPDKTILESPTD-KKVGWKVIFNNMNVNQNMGCPYDRDSWNPVYQ 194
Db 158 LSGKSFSTINYNKQESYRTTIDRTNHNKSGVGEAHKIMNMGWGPYGRDSYDPTYGNE 217

QY 255 DYQLHWTSTNWKGTNTKDKWDRSSERYKIDWE 287
Db 277 RYTQWNLRLHWGNKYNKQNTVTFSTVEVDWQ 309

RESULT 7
Q99T54
ID Q99T54 PRELIMINARY; PRT; 327 AA.
AC Q99T54;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Leukotoxin, Lkd.
GN Lkd OR SAl637.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J.-Q., Ito T.,
RA Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C.,
RA Hiraoka H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Ogasawara A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M.,
RA "Whole genome sequencing of meticillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003135; BAB42905.1;
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR Pfam; PF01117; Aerolysin; l.
DR PRINTS; PR01458; BICOMPENTOXIN.
DR TIGRFAMs; TIGR01002; hlyII; 1.
KW Complete proteome.
SQ SEQUENCE 327 AA; 36903 MW; COA5486421ACDAB9 CRC64;

Query Match      26.6%; Score 394.5; DB 16; Length 327;
Best Local Similarity 30.0%; Pred. No. 1.9e-15;
Matches 82; Conservative 70; Mismatches 112; Indels 9; Gaps 6;

QY 18 TTVKTGDLVTDKENGMMHKVFYSFDDKNHKKLLVIRTKGTIAGQYRVYSEEGANKSG 77
Db 43 TLYKTTATSDNDKLN-ISOILTFNFIKDKSYDKDTLVLAAGNINSYGKKPNKDYNSQ 101

QY 78 LAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEXMXTYTXGXNXXTXDXTKXGXKXG 137
Db 102 FYWGGKYNVSVSESNDVAVVYAPKN--QNEEFQVQOTLGSYG--GDINISNGLSGG 157

QY 138 XN--XSXGXTXKXVQDPDKTILESPTD-KKVGWKVIFNNMNVNQNMGCPYDRDSWNPVYQ 194
Db 158 LSGKSFSTINYNKQESYRTTIDRTNHNKSGVGEAHKIMNMGWGPYGRDSYDPTYGNE 217

QY 255 DYQLHWTSTNWKGTNTKDKWDRSSERYKIDWE 287
Db 277 RYTQWNLRLHWGNKYNKQNTVTFSTVEVDWQ 309
```

```
Db 102 FYWGGKYNVSVSESNDVAVVYAPKN--QNEEFQVQOTLGSYG--GDINISNGLSGG 157

QY 138 XN--XSXGXTXKXVQDPDKTILESPTD-KKVGWKVIFNNMNVNQNMGCPYDRDSWNPVYQ 194
Db 158 LSGKSFSTINYNKQESYRTTIDRTNHNKSGVGEAHKIMNMGWGPYGRDSYDPTYGNE 217

QY 195 LFMKTRNGSMKAADNFDLPNKASSLLSGSFDPFATVITMDRKASKQQTNDIVYERVD 254
Db 218 LFLGGRQSSNAGQNFPLTHQMPLLARGNFPNPEFISLVLSHKQNDTK-KSKIKVTYQREMD 276

QY 255 DYQLHWTSTNWKGTNTKDKWDRSSERYKIDWE 287
Db 277 RYTQWNLRLHWGNKYNKQNTVTFSTVEVDWQ 309

RESULT 8
Q931N4
ID Q931N4 PRELIMINARY; PRT; 332 AA.
AC Q931N4;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Leukotoxin F-subunit.
GN Lkd OR SAV1819.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Oi Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hiraoka H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RA "Whole genome sequencing of meticillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003363; BAB57981.1;
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR Pfam; PF01117; Aerolysin; l.
DR TIGRFAMs; TIGR01002; hlyII; 1.
KW Complete proteome.
SQ SEQUENCE 332 AA; 37738 MW; 5C896297453A3638 CRC64;

Query Match      26.6%; Score 394.5; DB 16; Length 332;
Best Local Similarity 30.0%; Pred. No. 1.9e-15;
Matches 82; Conservative 70; Mismatches 112; Indels 9; Gaps 6;

QY 18 TTVKTGDLVTDKENGMMHKVFYSFDDKNHKKLLVIRTKGTIAGQYRVYSEEGANKSG 77
Db 43 TLYKTTATSDNDKLN-ISOILTFNFIKDKSYDKDTLVLAAGNINSYGKKPNKDYNSQ 101

QY 78 LAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEXMXTYTXGXNXXTXDXTKXGXKXG 137
Db 102 FYWGGKYNVSVSESNDVAVVYAPKN--QNEEFQVQOTLGSYG--GDINISNGLSGG 157

QY 138 XN--XSXGXTXKXVQDPDKTILESPTD-KKVGWKVIFNNMNVNQNMGCPYDRDSWNPVYQ 194
Db 158 LSGKSFSTINYNKQESYRTTIDRTNHNKSGVGEAHKIMNMGWGPYGRDSYDPTYGNE 217

QY 195 LFMKTRNGSMKAADNFDLPNKASSLLSGSFDPFATVITMDRKASKQQTNDIVYERVD 254
Db 218 LFLGGRQSSNAGQNFPLTHQMPLLARGNFPNPEFISLVLSHKQNDTK-KSKIKVTYQREMD 276

QY 255 DYQLHWTSTNWKGTNTKDKWDRSSERYKIDWE 287
Db 277 RYTQWNLRLHWGNKYNKQNTVTFSTVEVDWQ 309
```

```
RESULT 9
Q53747
ID Q53747 PRELIMINARY; PRT; 325 AA.
AC Q53747;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE Leucocidin F precursor.
GN LUKF
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE OF 1-285 FROM N.A.
RC STRAIN=ATCC 49775;
RX Prevost G., Supersac G., Piemont Y., Colin D., Sire S., Henslen T.,
RA Petiau P., Meunier O., Gibier B., Koenig W., Piemont Y.;
RT "The new family of leukotoxins from S.aureus: structural and
biological properties.";
RL Zentralbl. Bakteriol. 24:284-293(1994).
[2]
RN SEQUENCE FROM N.A.
RP
RC STRAIN=ATCC 49775;
RA Prevost G.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP
RC STRAIN=ATCC 49775;
RX MEDLINE=96009775; PubMed=7558328;
RA Montell H., Couppie P., Petiau P., Supersac G., Finck-Barbancon V.,
RT "Panton-valentine leucocidin and gamma-hemolysin from
staphylococcus aureu ATCC 49775 are encoded by distinct loci and have
different biological activities.";
RL Infect. Immun. 63:4121-4129(1995).
[4]
RN EMBL; X72700; CAA51252.1;
DR InterPro: IPR001776; Aerolysin.
DR InterPro: IPR001340; Hemolysin_pore.
DR Pfam: PF01117; Aerolysin; 1.
DR PRINTS: PR01468; BICOMPOTOXIN.
DR TIGRFAMs: TIGR01002; hlyII; 1.
KW SIGNAL.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 325 POTENTIAL.
SQ SEQUENCE 325 AA; 37009 MW; 03C2C247B3B3CCD9 CRC64;

Query Match 26.5%; Score 394; DB 2; Length 325;
Best Local Similarity 28.6%; Pred. No. 2e-15;
Matches 80; Conservative 71; Mismatches 109; Indels 20; Gaps 6;

QY 11 TTDIGSTVTKGDLVTVYDKENGHHKVPYSFIDRNNHKKLLVIRTKGTAGQYRVSE 70
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 45 TTATSDSLKISKILT-----FNFIKDSYDKDTLILKAAGNIYSGYTKPNP 92
QY 71 EGANKSGLAWPSAFKVLQLPDNEVAQISDYIPRNSIDTXEXMXTXGKNNXNTXDTX 130
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 93 KDTISQFTWGSYKINISINDSNDVNVVDYAPKN--QNEEFQVQQTGVISYG--GDINI 148
QY 131 KXGLXGXN--XSXGXTXKXVPDFKXLTILESPTD-KKVGKVIFFNNVNNWGPYDRDSW 187
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 SNGLSGGNGSKSFSETINKQESYRTSLDKRNFNFKIGWDVEAHKIMNNGWGPYGRDSY 208
QY 188 NPVYGQLFMKTRNGSKAADFLDPNKAASSLLSSGSPDFATVTPMDRKASKQQTNDIV 247
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 HSTYGNMFLGSRQSLNAGQNFLEYHKPVLRSRGNFEPFGLVSKRQNAAK-KSKITV 267
QY 248 IYERVRDDYQLHWTSTNWKGTNTKDKWTRDSERYKIDWE 287
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 TYQSEMDRYTNFNFNWNIGNKKYKHIRATHTSIYEDWE 307
```

```
RESULT 10
Q57227
ID Q57227 PRELIMINARY; PRT; 325 AA.
AC Q57227;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE HLBG-like ORF precursor.
GN HLBG-LIKE ORF OR LUK F-R.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 49775;
RX MEDLINE=96009775; PubMed=7558328;
RA Prevost G., Couppie P., Petiau P., Supersac G., Finck-Barbancon V.,
RT "Panton-valentine leucocidin and gamma-hemolysin from Staphylococcus
aureus ATCC 49775 are encoded by distinct genetic loci and have
different biological activities.";
RL Infect. Immun. 63:4121-4129(1995).
[2]
RN SEQUENCE FROM N.A.
RP
RC STRAIN=P83;
RA Prevost G.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE OF 1-186 FROM N.A.
RP
RC STRAIN=P83;
RX MEDLINE=93138781; PubMed=8423088;
RA Supersac G., Prevost G., Piemont Y.;
RT "sequencing of leucocidin R from Staphylococcus aureus P83 suggests
that staphylococcal leucocidins and gamma-hemolysin are members of a
single, two-component family of toxins.";
RL Infect. Immun. 61:580-587(1993).
[4]
RN SEQUENCE FROM N.A.
RP
RC STRAIN=P83;
RA Prevost G.;
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
[5]
RN SEQUENCE FROM N.A.
RP
RC STRAIN=P83;
RA Prevost G.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
[6]
RN SEQUENCE FROM N.A.
RP
RC STRAIN=P83;
RX MEDLINE=95210788; PubMed=7696748;
RA Prevost G., Supersac G., Piemont Y., Colin D., Sire S., Henslen T.,
RA Petiau P., Meunier O., Gibier B., Koenig W., Piemont Y.;
RT "A particular class of virulence factors: calmodulin-activated
bacterial adenylate cyclases.";
RL Zentralbl. Bakteriol. 24:284-293(1994).
DR EMBL; X81586; CAA57278.1;
DR EMBL; X64389; CAA45729.1;
DR HSSP; P09616; 7AHL.
DR InterPro: IPR001776; Aerolysin.
DR InterPro: IPR001340; Hemolysin_pore.
DR InterPro: IPR003963; Staph_bicn_tkn.
DR Pfam: PF01117; Aerolysin; 1.
DR PRINTS: PR01468; BICOMPOTOXIN.
DR TIGRFAMs: TIGR01002; hlyII; 1.
KW SIGNAL.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 325
SQ SEQUENCE 325 AA; 36812 MW; 38A0066D8DBD8935 CRC64;

Query Match 26.5%; Score 393; DB 2; Length 325;
```

Best Local Similarity 29.2%; Pred. No. 2.3e-15;
Matches 83; Conservative 72; Mismatches 109; Indels 20; Gaps 6;

```
Qy 11 TTDTGNTTAVKGLVTDKENGMMHKVFYFIDDKNNHKLVLVTKGTIAGQYRVYSE 70
Db 47 TTATADSKFKISQILT-----FNIKDKSDKDTFLVKAAGNINGSYERPNP 94
Qy 71 EGANKSGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEMXT--XTXGXNXXNTDX 128
Db 95 KDYDFSKIYWGAKYNNVSISSQSDSNVVDYAPKQNEEFQVONTLGYTFGGDISNGL 154
Qy 129 TKXGXKXNXXGXTKXXVQDPDKTILESPTD-KKVGWKVIFNNMVNQNNQYDRDSW 187
Db 155 ---SGGLAG-NTAFSETINYKQESYRTTILSRNTNKNVCGWVEAHKIMNNGWGPYGRDSF 210
Qy 188 NPVYGNQLFMKTRNGSKAADNFDLPNKASSLLSGSPDFATVITMDRKASKQQTNDV 247
Db 211 HPTYNELFLAGRSQSAVAGQNFIAQHOMPLLSRSNFNPEFLSVLS-HRQDGAKKSKITV 269
Qy 248 IYERVDDYQLHWTSTNKGTTKDKWTDSSERYKIDWEKEEM 291
Db 270 TYQREMDLYQIRWNGFYWAGANYKNFKTRTFKSTYEIDWENHKV 313
```

```
RESULT 11
Q931F3 PRELIMINARY; PRT; 325 AA.
AC Q931F3;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Gamma-hemolysin component B.
GN HLCB OR SAV2421.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003365; BAB58583.1;
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR Pfam; PF01117; Aerolysin; 1.
DR TIGRFAMs; TIGR01002; hlyII; 1.
KW Complete proteome.
SQ SEQUENCE 325 AA; 36658 MW; 082999EB877CA2B5 CRC64;
```

Query Match 25.9%; Score 385; DB 16; Length 325;
Best Local Similarity 29.9%; Pred. No. 6.7e-15;
Matches 89; Conservative 73; Mismatches 116; Indels 20; Gaps 8;

```
Qy 9 TGTDTI-GSNTTV---KTGDLVTDKENG-----MHKKVFYFIDDKNNHKLVLVIR 56
Db 21 SGTANAEGKITPVSKKVDVDTLYKITATADSKDKFISQILTFFNFKDKSYDKDTLVK 80
Qy 57 TGTGTAGYRVYSEGANGLAWPSAFKVLQLPDNEVAQISDYPRNSIDTXEMXT- 115
Db 81 ATGAINSGFVAPNPNDYDFSKLYWGAKYNVSISSQSDSNVVDYAPKQNEEFQVQNTL 140
Qy 116 -XTGXNXXNTDXTKXGXKXNXXGXTKXXVQDPDKTILESPTD-KKVGWKVIFNN 173
Db 141 GTTFGGDISNGL---SGGLNG-NTAFSETINYKQESYRTTILSRNTNKNVCGWVEAHK 196
```

```
Qy 174 MVNQNGPYDRDSWNPVYGNQLFMKTRNGSKAADNFDLPNKASSLLSGSPDFATVIT 233
Db 197 IMNNGWGPYGRDSFPTTYGNELFLAGRSQSAVAGQNFIAQHOMPLLSRSNFNPEFLSVLS 256
Qy 234 MDRKASKQQTNDVYERVDDYQLHWTSTNKGTTKDKWTDSSERYKIDWEKEEM 291
Db 257 -HRQDGAKKSKITVTYQREMDLYQIRWNGFYWAGANYKNFKTRTFKSTYEIDWENHKV 313
```

RESULT 12

```
Q99SN8 PRELIMINARY; PRT; 338 AA.
AC Q99SN8;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Hypothetical protein SAV2004.
GN SAV2004 OR SA1812.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX SPECIES=S. aureus (strain Mu50), and S. aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003364; BAB58166.1;
DR EMBL; AP003335; BAB43092.1;
DR HSSP; P09616; 7AHL.
DR InterPro; IPR001776; Aerolysin.
DR InterPro; IPR001340; Hemlysn_pore.
DR InterPro; IPR003963; Staph_bicn_txn.
DR Pfam; PF01117; Aerolysin; 1.
DR PRINTS; PR01468; BICOMPOTOXIN.
DR TIGRFAMs; TIGR01002; hlyII; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 338 AA; 38662 MW; B4B4F335B4B8DB1 CRC64;
```

Query Match 25.9%; Score 385; DB 16; Length 338;
Best Local Similarity 28.2%; Pred. No. 7.2e-15;
Matches 82; Conservative 81; Mismatches 110; Indels 18; Gaps 8;

```
Qy 13 DIGSNTTVKGLVTDKENGMMHKVFYFIDDKNNHKLVLVTKGTIAGYRVYSEEG 72
Db 42 NLGDGTYMYRTATTSQSKNITQSLQFNFLTEPNYDKETVFIKAKGTIGSLRIDPNG 101
Qy 73 ANKSGLAWSAFKVLQLPDNEVAQISDYPRNSIDTXEMXTYTXGXNXXNTDXTKX 131
Db 102 YWNTSLRWPGSYSVSIQNVDDNNNTNTVDFAPKQDESREVKYT----YGVKTGDFSN 157
Qy 132 XGXKLG---XNKSXGXTKXXVQDPDKTILESPTDKK-VGVKVV---IFNNMVNQNGPYDR 184
Db 158 RGLTGNTKRESNSETISYQPSYRFLLDQSTSHKGVGWKVEAHLNNMGHDHTRQLTN 217
Qy 185 DSWNPVYGNQLFMKTRNGSKAADNFDLPNKASSLLSGSPDFATVITMDRKASKQQTN 244
Db 218 DSDNRT-KSEIFSLTRNGNLWAKNFTPKDKMPVTVSEGFNPEFLVAMSHDKK-DKGSQ 275
Qy 245 IDVYERVDDYQLHWTSTN---WKGNTNTKDKWTDSSERYKIDWEKEEM 291
Db 276 FVHYKRSMDEFKIDWNRHGFNGWYSGENHVDKKEELKSLALYEVDWKTHDV 326
```



```
DR Pfam: PF01117; Aerolysin; 1.  
DR TIGRFAMs: TIGR01002; hlyII; 1.  
SQ SEQUENCE 322 AA; 36519 MW; 9424E58715B108E6 CRC64;  
  
Query Match 25.28; Score 374; DB 9; Length 322;  
Best Local Similarity 27.5%; Pred. No. 3e-14;  
Matches 84; Conservative 75; Mismatches 119; Indels 28; Gaps 6;  
  
Qy 6 NIKGTGTTDIGSNTVKTGDLVYDKENGMHKKV-----FYSFIDD 45  
Db 10 SVATSIILMLNSNTVDAQAHTVPSEKKVDDKITLYKTATSDSKLKISQILTFNFIKD 69  
  
Qy 46 KNHNKKLIVIRTKGTIAGQYRVYSEEGANKSGLAWPSAFKVLQLPDPNEVAQISDIYPRN 105  
Db 70 KSYDKDTLILKAAGNIYSGYTQPTSDSSINSQFYWGAKYNVFVSSEKDSYNIIVDYPKN 129  
  
Qy 106 SIDTXEXMXT--XYGXNXXNTXDXTKXGXNXXSGXTXKXVQPDFKXTILESPTD- 162  
Db 130 QNEEFVQOQTGLGYSYGGDININGLT---GGLNG--SKSFSETINYKQESYRTTIDRKTNH 185  
  
Qy 163 KKVGVKVIENMVNQNMGPDYDRDSWNPVYGNLFEMKTEGSMKAADNFDLPNKASSLLSS 222  
Db 186 KSIQGVVAHKIMNMGWGPYGRDSDSLYGNELFGLGQSSNANQNPLPTHQMPILARG 245  
  
Qy 223 GFSPDFATVITMDRKASKQQTNIIDVIYERVRDDYQLHWTSTNWKGTNTKDKWTDSSERY 282  
Db 246 NFNPEFISVLSHKQKDYK-KSKIKVTYQRENDRYENFWNNLHWIGYNIKNQKRATHTSIY 304  
  
Qy 283 KIDWEK 288  
Db 305 EIDWEK 310
```

Search completed: March 4, 2003, 10:44:08
Job time : 30.5 secs